MASTER THESIS

Course code: BE305E

Candidate name: Pavel Serin

Is random walk hypothesis a reasonable data generating process assumption for stock prices?

Date: 19.05.2017

Total number of pages: 45

NORD University

www.nord.no

Sammendrag

Denne masteroppgaven undersøker kurssvingninger på aksjeindekser. Hovedmålet med avhandlingen er å teste hvor random walk hypotesen en rimelig datagenererende prosessen forutsetning for aksjekurser? Den viktigste motivasjonen for studien er å få bedre forståelse av naturen av prisendringer på børsen. Det er en teoretisk gjennomgang av eksisterende studier på dette området. Jeg vurdere ulike synspunkter fra ulike forskere som gir fakta som støtter begge sider av spørsmålet. Jeg beskriver også metode for studien og gjennomføre ulike tester som sjekker hypotesen om random walk. Men, gjør resultatene av studien ikke gi eksakte svar på stilte spørsmålet gitt i emnet. Fremtidige undersøkelser i dette området kan være nødvendig.

Table of contents

1. Introduction

The minds of scientists have been occupied for a very long time with the question of how prices are generated on the stock market. Unfortunately, it is impossible to get a definite answer on this question. However, lots of researches conducted numerous tests that evaluated the degree of randomness or predictability of market movements. In this work I've attempted to perform an overview of different pricing models (random walk, mixture model/Markov switching model) and their theoretical justification.

Investors have different opinions on the same issue. Some statisticians and econometricians have argued that price movements are impossible to predict, since it is subject to a random movement, and repetitive schemes and patterns of no more than chance (Fama, 1995). According to this view it is impossible to exploit vulnerabilities or inefficiencies in the market for earning more than expected profits. Others have stated that from time to time there appear chances of using techniques that allow to predict the future price change with a certain degree of probability. The forecast could be based on a variety of things, for example, some proxy variables which allow to predict the risk premium of the assets basing the decision on the levels of asset prices (Keim & Stambaugh, 1986). A history of previous price changes of securities could be used. Or pieces of information about a fundamental intrinsic value also may be useful (Abarbanell & Bushee, 1997).

Technical analysts believe that all the information about the asset enclosed in the history of previous prices. They convinced that history repeats itself, so if one recognizes patterns of price behavior, it can be used for earning extraordinary profits, more than average investor earns. Fundamentalists are trying to find some internal basic value of the asset and compare it with the market. Then they build their strategies on the basis of the information received. However, if the random walk of stock prices actually takes place in reality, then all this story about history and fundamental price is not worth a penny.

This paper is organized as follows: in the first part I am having a look on the definition of a time series concept, random walk and its characteristics. Then we review examples of different price modelling in the literature, e.g. random walk model, constant expected return model and Markov switching model. In part of literature review we are having a brief look on the previous studies of that topic. Later we discuss the methodology used for hypothesis testing, tests listing and their explanation. Results and conclusions are presented in the final part of the paper.

1.1. Reviewing the time series

Analysis of past statistical data has always been one of the main methods for predicting the behavior of some phenomenon. In particular, in econometrics mathematical models which are based on empirical data are used for prediction of economic processes. Time series analysis is no exception, but its investigation appears to be somewhat more complicated rather than simple crosssectional data due to various problems that might occur. These are the possible autocorrelation of residuals, non-stationarity of the time series, seasonal dependence and many other problems. However, there are tools and methods which can help solving these problems.

A time series is a sequence of measured through some (usually equal) intervals of time data. Time series analysis combines methods for studying time series as trying to understand the nature of the data points. In particular, it tries to give answers on questions like "what caused the variable to behave like that?", "are there any interdependences between one and other variables?". It is also attempting to build a forecast for the future. Prediction of time series is based on the model construction which gives out possible future events basing the forecast on the previous data. A typical example is the opening price prediction on the stock exchange based on previous trading activities.

The market prices of stocks, bonds and other securities are typical example of time series. Moreover, their changes and attempts to predict their behavior is the task of an army of analysts and traders worldwide. Price fluctuations in the stock market can have an impact on the macroeconomic situation in a given country (Asprem, 1989), and in the world as a whole (Beber & Pagano, 2013). Therefore, this paper will try to analyze available data and reduce the degree of uncertainty in this matter. That is why an understanding of the possible change in the yield of securities is an actual problem nowadays both in science and in business.

1.2. Random walk as autoregressive model of order one

In order to lay the foundations for hypothesis testing that will be performed in this Master Thesis seems reasonable to clarify what is meant by the term "random walk" itself. A time series is called stationary if three conditions of stationarity (weak) are satisfied:

1. The mean of the distribution is independent of time.
2. The variance of the distribution is independent of time.
3. The covariance between its values at any two time points depends only on the distance between those points, and not on time.

For the time series to be stationary, the β_2 coefficient in the model:

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

should have absolute value less than 1. It is easy to show that in this case all three conditions of weak stationarity are satisfied.

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

$$X_{t-1} = \beta_2 X_{t-2} + \varepsilon_{t-1}$$

$$X_t = \beta_2^2 X_{t-2} + \beta_2 \varepsilon_{t-1} + \varepsilon_t$$

$$X_t = \beta_2^t X_0 + \beta_2^{t-1} \varepsilon_t + \cdots + \beta_2^2 \varepsilon_{t-2} + \beta_2 \varepsilon_{t-1} + \varepsilon_t$$

$$E(X_t) = \beta_2^t X_0$$

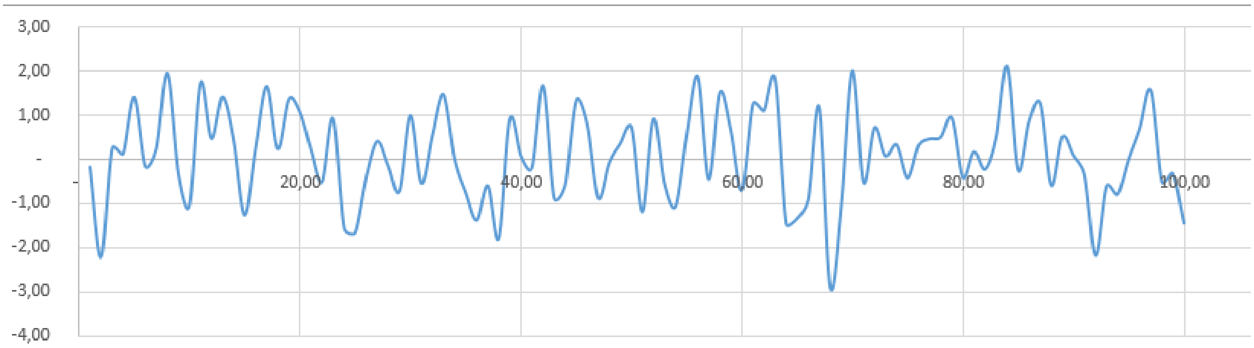
Hence E(X_t) = $\beta_2^t X_0$ since the expected value of each new innovation is zero. Since the expectation is not a function of time, the first condition is satisfied.

$$var(X_t) = (\frac{1-\beta_2^{2t}}{1-\beta_2^2})\sigma_\varepsilon^2 \rightarrow (\frac{1}{1-\beta_2^2})\sigma_\varepsilon^2$$

If the absolute value of $\beta_2 < 1$, β_t^2 tends to zero as t increases. Thus, ignoring transitory (short lived) initial effects, the variance tends to a limit that is independent of time.

$$cov(X_t, X_{t+s}) = cov(X_t, \beta_2^s X_t) + cov(X_t, [\beta_2^{s-1}\varepsilon_{t+1} + \cdots + \beta_2^2\varepsilon_{t+s-2} + \beta_2\varepsilon_{t+s-1} + \varepsilon_{t+s}])$$
$$= \beta_2^s var(X_t)$$

As we have just seen, var(X_t) is independent of t, apart from a transitory initial effect. Hence the third condition for stationarity is also satisfied.

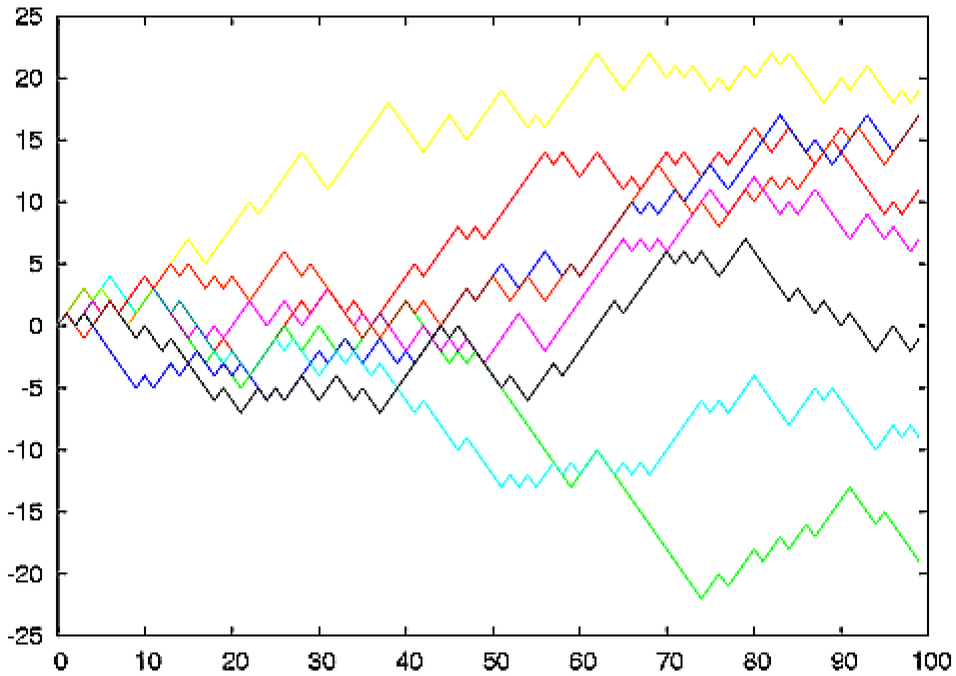


Picture 1 Example of stationary time series

As we know (see for example Dougherty (2007) the random walk is a nonstationary time series in which the second condition of weak stationarity is violated. Consider the following autoregressive model of order 1:

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

Where β_2 is equal to 1. In words, the value of X in time period t is equal to its value in the time period t-1, plus a random adjustment. This is what we call random walk. In this case variance of the process is proportional to t, which violates stationarity.



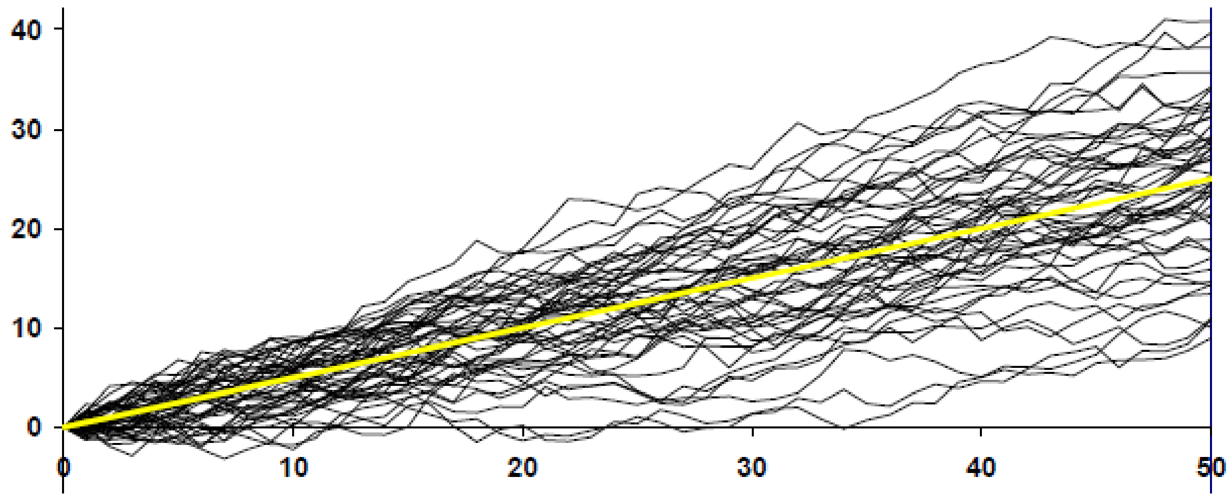
Picture 2 Example of random walk ensemble distribution

One possible modification of this type of time series is so called random walk with drift. It can be made by inserting a constant term into equation:

$$X_t = \beta_1 + \beta_2 X_{t-1} + \varepsilon_t$$

The difference lies in the fact that now mean of the process becomes a function of time, which violates the first condition for stationarity:

$$E(X_t) = t\beta_1$$



Picture 3 Example of random walk with drift ensemble distribution

1.3. Constant expected return as static model

Constant expected return (CER) model assumes that returns of the assets are i.i.d. (independently and identically distributed) and have constant mean and variance. It is possible for this model to have contemporaneously correlated returns on different assets however they have to be constant over time. The CER model is widely used in finance and econometrics. For example, it is used in mean-variance portfolio analysis, the Capital Asset Pricing model (CAPM), and the Black-Scholes option pricing model. Although this model is very simple, it provides important intuition about the statistical behavior of asset returns and prices and serves as a benchmark against which more complicated models can be compared and evaluated. It allows the discussion and development of several important econometric topics such as Monte Carlo simulation, estimation, bootstrapping, hypothesis testing, forecasting and model evaluation.

In CER the continuously compounded return on asset I at time t denoted as $r_{it} = \ln(\frac{P_{it}}{P_{t-1}})$. There are several assumptions regarding the probability distribution of r_{it} for i = 1,…,N assets over time t = 1,…,T:

1) Covariance stationarity and ergodicity: $\{r_{i1}, \dots, r_{iT}\} = \{r_{it}\}_{t=1}^{T}$ is a covariance stationary and ergodic stochastic process with$E[r_{it}] = \mu_i, var(r_{it}) = \sigma_i^2$, $cov(r_{it}, r_{jt}) = \sigma_{ij}$ and $cor(r_{it}, r_{jt}) = \rho_{ij}$
2) Normality: $r_{it} \sim N(\mu_i, \sigma_i^2)$ for all i and t.
3) No serial correlation: $cov(r_{it}, r_{js}) = cor(r_{it}, r_{is}) = 0$ for t ≠ s and i,j = 1,…,N.

Those assumptions state that in each time period returns are normally distributed. Means and variances, covariances and correlations between assets are constant. Assets returns are serially uncorrelated

$cov(r_{it}, r_{js}) = cor(r_{it}, r_{is}) = 0$ for all i and t ≠ s

and the returns on all possible pairs of assets i and j are serially uncorrelated

$cov(r_{it}, r_{js}) = cor(r_{it}, r_{js}) = 0$ for all i ≠ j and t ≠ s

It is obvious that those assumptions are very strong and partly unrealistic. However, they allow the development of straightforward probabilistic model for asset returns as well as statistical tools for estimating the parameters of the model, testing hypotheses about the parameter values and assumptions.

Traditional CER regression model looks like this:

$$r_{it} = \mu_i + \varepsilon_{it}$$

where ε_{it} is a Gaussian white noise process with zero expectation ($E[\varepsilon_{it}] = 0$) and variance $= \sigma_i^2$.

The Constant Expected Return obtains very simple form and claims that each asset return is equal to a particular constant μ_i which reflects the expected return plus a normally distributed random variable ε_{it} with mean zero and constant variance. This random component can be interpreted as unexpected news concerning the value of the asset that arrives between to points in time, t-1 and t. It implies that

$$\varepsilon_{it} = r_{it} - \mu_i = r_{it} - E[r_{it}]$$

So that ε_{it} is a deviation of assets return from its expected value. In case of good news between time points t-1 and t, the realized value of ε_{it} is positive and total return is higher than predicted by the model. And vice versa, if the news is bad then the return will be lower than expected. In the long run all deviations from the expected value should give zero sum implying that on average news are neutral, neither good nor bad. The assumption that variance $\varepsilon_{it} = \sigma_i^2$ can be interpreted as saying that volatility, or typical magnitude, of news arrival is constant over time.

The CER model of asset returns gives rise to the so-called random walk model for the logarithm of asset prices. Letting $p_{it} = \ln(P_{it})$ and using the representation of r_{it} in the CER model, it is possible to express the log-price as:

$$p_{it} = p_{it-1} + \mu_i + \varepsilon_{it}$$

This representation is known as random walk model in log-prices. The RW model provides the following interpretation for the evolutionary process of log prices. Let p_{i0} denote the initial log price of asset i. The RW model claims that the log-price at time t = 1 is

$$p_{i1} = p_{i0} + \mu_i + \varepsilon_{i1}$$

By repeated recursive substitution, the log price at time t = T is

$$p_{iT} = p_{i0} + T * \mu_i + \sum_{t=1}^{T} \varepsilon_{it}$$

The actual price, p_{iT}, deviates from the expected price by the accumulated random news:

$$p_{iT} - E[p_{iT}] = \sum_{t=1}^{T} \varepsilon_{it}$$

At time t = 0, the variance of the log-price at time T is:

$$var(p_{iT}) = var\left(\sum_{t=1}^{T} \varepsilon_{it}\right) = T * \sigma_i^2$$

Hence, the RW model implies that the random process of log-prices $[p_{it}]$ is non-stationary because the variance of p_{it} rises with the flow of time t.

1.4. Markov switching model

The Markov switching (or regime switching) model is a popular nonlinear time series model. It involves multiple equations that characterize behavior of time series in various regimes. This model can capture complex dynamic parameters due to permission of switching between equation. As Kuan (2002) claims in his article: “A novel feature of the Markov switching model is that the switching mechanism is controlled by an unobservable state variable that follows a first-order Markov chain. In particular, the Markovian property regulates that the current value of the state variable depends on its immediate past value. As such, a structure may prevail for a random period of time, and it will be replaced by another structure when a switching takes place".

Various evidences suggest that TS behavior of financial and economic variables follow multiple patterns over time. So instead of using single model for the conditional mean of a variable, seems logical to employ several models to represent these patterns. A Markov switching model is built by mixing two or more dynamic models via a Markovian switching mechanism.

The simple Markov switching model looks like this. Let s_t denote an unobservable state variable assuming the value 1 or 0. A simple switching model for the variable z_t involves two autoregressive specifications:

$$z_t = \begin{cases} \alpha_0 + \beta z_{t-1} + \varepsilon_t, & s_t = 0 \\ \alpha_0 + \alpha_1 + \beta z_{t-1} + \varepsilon_t, & s_t = 1 \end{cases}$$

where β has absolute value less than 1 and the error term ε_t is i.i.d. random variables with mean zero and variance σ_ε^2.

This is a stationary autoregressive process of order 1 with mean = $\alpha_0/(1-\beta)$ when s_t = 0, and it switches to another stationary AR(1) process with mean = $(\alpha_0 + \alpha_1)/(1-\beta)$ when s_t switches from 0 to 1. Then provided that $\alpha_1 \neq 0$, this model admits two dynamic structures at different levels, depending on the value of the state variable s_t. In this case, z_t are controlled by two different distributions with distinct means, and s_t determines the switching between these two regimes.

When $s_t = 0$ for $t = 1,\ldots,\tau_0$ and $s_t = 1$ for $t = \tau_0 + 1,\ldots,T$, the previously mentioned model is the model with a single structural change in which the model parameter experiences only one change after $t = \tau_0$. In the random switching model, the realization of s_t is independent of the previous and future states so that z_t may switch back and forth between different states. If s_t is postulated as the indicator variable $1_{[\lambda t \leq c]}$ such that s_t = 0 or 1 depending on whether the value of λ_t is greater than the cut-off value c, previous model becomes a threshold model. It is quite common to choose a lagged dependent variable, e.g. z_{t-d} as the variable λ_t.

Although these models can describe the behavior of the time series in the two states, both have limitations. Only one admitted change is very little amount. The straightforward solution is extending the model allowing multiple changes. But typically the results are bulky and unsatisfactory (Bai, 1999; Bai & Perron, 1998). Also the time is exogenous in such models but it determines changes in them. By contrast random switching model allows several changes, however with state variables exogenous to the dynamic structures in the model. "This model also suffers from the drawback that the state variables are independent over time and hence may not be applicable to time series data. On the other hand, switching in the threshold model is dependent and endogenous and results in multiple changes. Choosing a suitable variable λ_t and the threshold value c for this model is usually a difficult task, however" (Kuan, 2002).

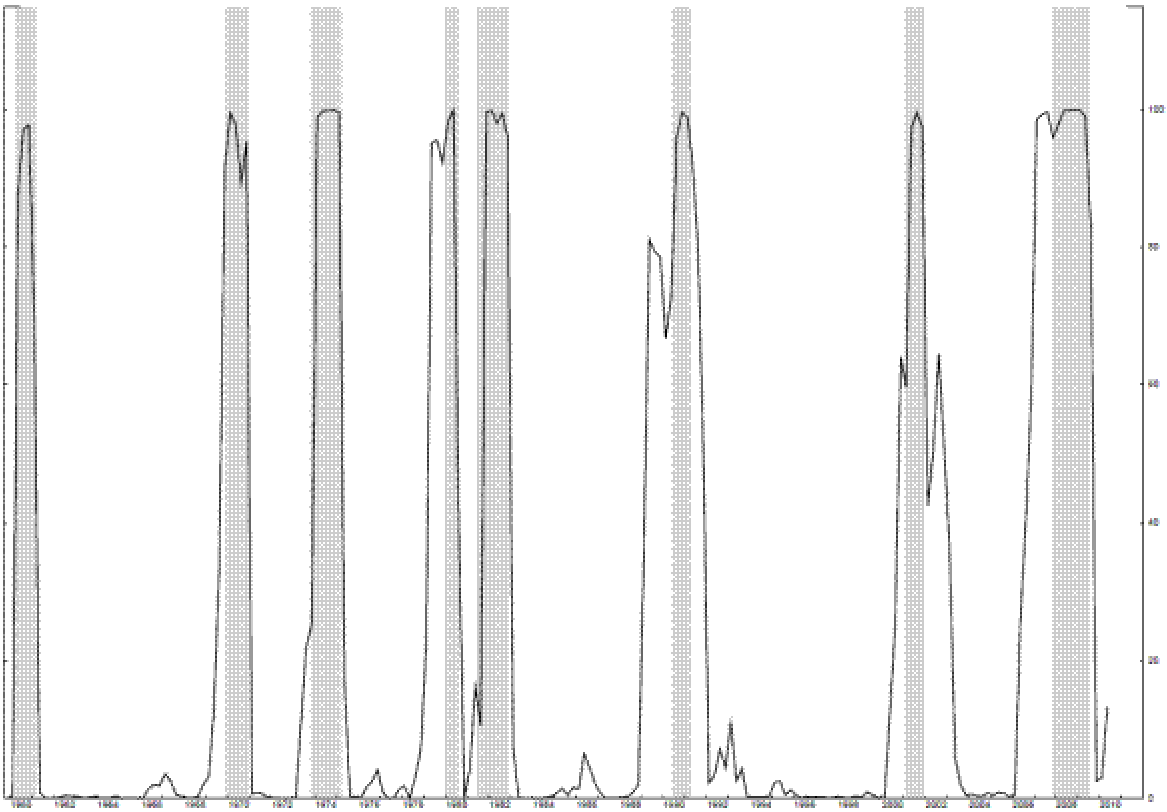
One way of solving mentioned problems is considering different specification for s_t. For example, consider s_t following first order Markov chain with the transition matrix presented below:

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

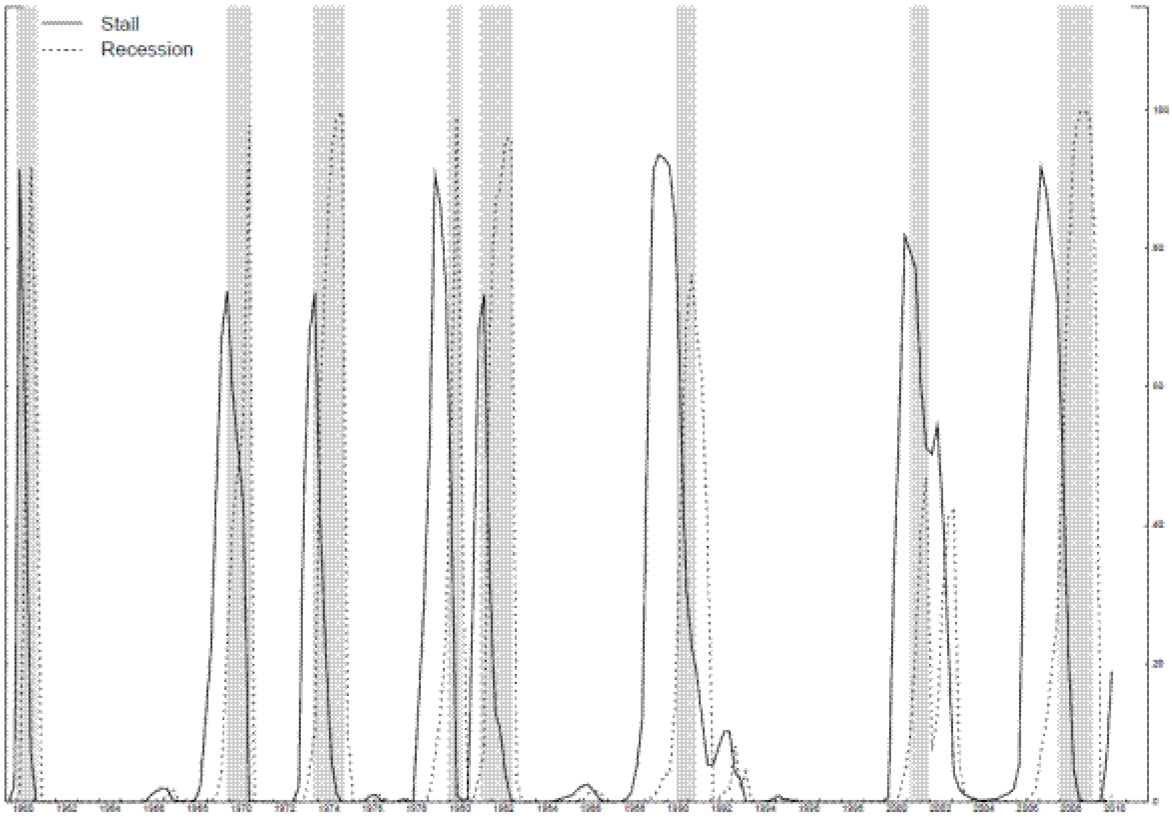
where p_{ij} (i, j = 0, 1) denote the transition probabilities of $s_t = j$ given that $s_{t-1} = i$. Obviously, the transition probabilities satisfy $p_{i0}+p_{i1} = 1$. This transition matrix controls the random behavior of the state variable. It contains only two parameters, p_{00} and p_{11}.

In the Markov switching model, the characteristics of z_t are jointly defined by the random properties of the driving innovations ε_t and the state variable s_t. In particular, the Markovian state variable yields random and frequent changes of model structures. Its probability of transition determines the persistence of each regime. A difficulty with the Markov switching model is that it may be tough to interpret because the state variables are unobservable.

Pictures below illustrate two and three state Markov Switching Models (Nalewaik, 2011):



Picture 4 Smoothed probabilities of Low Growth State, After BEA's3rd 2010 Q2 data release, GDP and GDI



Picture 5 Smoothed Probabilities, After BEA's 3rd 2010 Q2 data release, GDP and GDI

2. Literature review

2.1. Efficient market hypothesis

The main purpose of capital market is to provide access to financial resources and "moving" it from areas of its excess into the areas of its deficit. This process is something similar to movement of air masses in the atmosphere. Therefore, an absolutely effective market is a market in which asset prices will be completely accurate and informative indicators of the information available to investors at a particular moment of time. The information is instantly reflected in the price. There are many formulations of what the term "effective market", but in general all definitions point out the same thing: an investor on the absolutely efficient markets cannot get excess returns over the regular amount by using the ineffectiveness of market, exploiting the "loopholes" in the market or searching for undervalued or overvalued stocks (Malkiel & Fama, 1970).

According to Fama (1970) market efficiency could be divided into three categories: weak form, semi strong form and strong form of market efficiency. In the presence of a weak market efficiency it is assumed that all information about an asset, stock or security is contained in the historical prices and no investor can extract extra profits focusing solely on past prices. Serial correlation is absent and market return adheres to a certain constant mean (Poshakwale, 1996). Form, which includes more information is called semi strong form. In this form of market efficiency, any public information available to investors is also reflected in the price (news about unexpected earnings, stock splits, dividends, IPOs, etc.). Strong form of market efficiency includes the two previous ones, as well as takes into account the private information which is not available to a wide range of investors.

It worth noting that efficient market hypothesis (EMH) does not claim that investors will never gain high excess return. On the contrary, they might be high and low, the main point is not here. The rule is that it's impossible to gain extra profits systematically. One cannot be lucky every time to find inefficiencies, so the expected value of excess return is zero. Even though the market information is open and public, traders interpret it in different ways, so because of this disagreement random walk will arise around certain average price. More and more traders will try to find a scheme that describes and predicts the behavior of prices and use it. However, the more people will exploit it, the less effective it will be. After all it comes to naught and random walk continues.

According to well-known specialist in stock market analysis, author of the classic finance book *"A Random Walk Down Wall Street"* Burton Malkiel, additional tools like technical and fundamental analysis fail to provide excess profits to the investor above the average market risk premium (Malkiel, 2003). Technical analysis, which is the study of historical securities prices in order to determine future prices and fundamental analysis, which is the search in the financial information available about the security, helps investors to find so-called "undervalued" stocks. However, they fail to do it and cannot provide returns greater than those that could be provided by randomly selected portfolio. Lots of strategies failed to beat buy-and-hold technic in long term run. So can we say that this is a proof of existing random walk on the stock market? Unfortunately not. However, the fact that buy-and-hold seemed to be a winning strategy, I cannot claim that this is a proof of random stock price movement. There are opposing studies which show that buy-and-hold outperforms other strategies (Spinu, 2015) and studies that there are technics and tools (e.g. Shiryaev-Zhou index) which perform higher returns than buy-and-hold (Hui & Yam, 2014).

One of the reason that prevent momentum traders to earn extra profits rather than simple buy-and-hold returns is transaction costs. Barber and Odean (1999) show the sample of investors that made far worse than long term buy-and-hold traders. What is more interesting is that it happened during the period of well observable positive trend which could be exploited. But high transactional costs eliminated all the attempts. Something similar was discovered by Lesmond, Schill, and Zhou (2004). They have evaluated the profitability of relative strength trading strategies which consisted in buying past highly profitable securities and selling past low profitable. As a result, high transactional costs destroyed all the profits of the strategy due to type and frequency of securities traded.

Charles and Darné (2009) study Chinese market in their article describing efficiency of the Chinese stock markets. Since the establishment of two exchange systems – Shanghai Stock Exchange and Shenzhen Stock Exchange – they expanded rapidly and operated in a continually developing regulatory environment. Nowadays China's stock market is the second largest in Asia, behind only Japan. One of the possible scenario is that China's securities market has the potential to rank among the top four or five in the world within the coming decade.

The study of daily data for the Shanghai and Shenzhen stock markets was chosen for two types of shares: A and B over the period 1992–2007. A shares are denominated and traded in the local currency while B shares are traded in foreign currency. Such a large sample provided authors with a greater amount of information and reflected the significant changes that had taken place in China's securities sector in that period. They also investigate the EMH over various sub-periods in order to analyze the effects of the important changes in the relationship between the banks and the stock market in 1996 and 2000 as well as those of the implementation of the new policy allowing.

The results revealed by Charles and Darné (2009) suggest that Class A shares appear more efficient than Class B shares. This means that liquidity, market capitalization and information asymmetry might play a role in providing the explanation of the weak-form efficiency. In the same time, B shares for Chinese stock exchanges do not follow the random walk hypothesis and therefore are significantly inefficient. However, they appear to be efficient after the re-entry of banks in the stock market. Sow the entry of Chinese investors to the B-class share market have made positive impact the B-share market efficiency. Further research should investigate the effects of the re-entry of banks and the entry of domestic investors on the efficiency of Class B shares.

In general, I can quote Malkiel (2003), who characterizes well market participants' attempts to use apparent ineffectiveness of the system: "The general problem with these predictable patterns or anomalies, however, is that they are not dependable from period to period. Wall Street traders often joke that now the "January effect" is more likely to occur on the previous Thanksgiving. Moreover, these non-random effects (even if they were dependable) are very small relative to the transactions costs involved in trying to exploit them. They do not appear to offer arbitrage opportunities that would enable investors to make excess risk-adjusted returns".

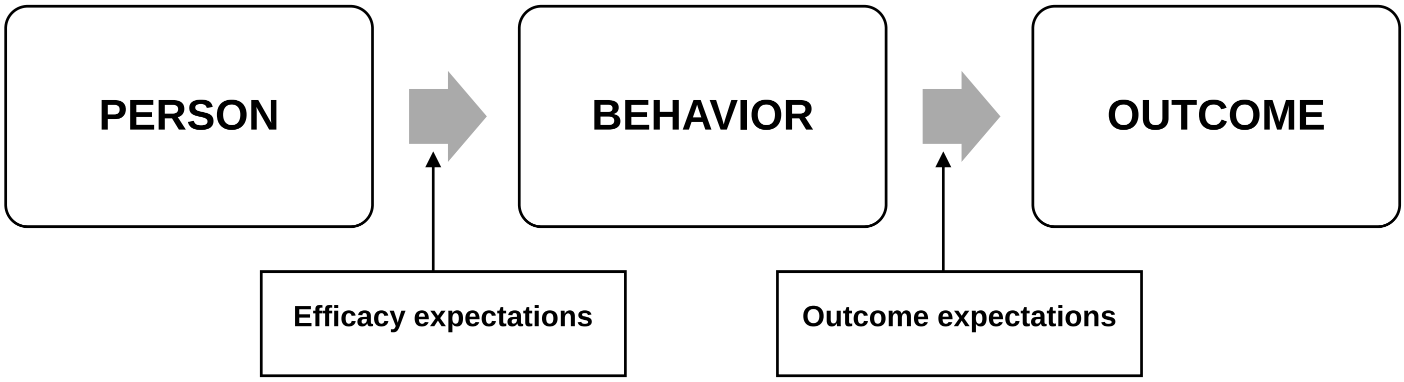
2.2. Social cognitive theory

Another candidate for the role of explainer of changes in securities prices is the social cognitive theory.

Self-efficacy can be defined as one's own confidence in capability to perform certain actions on the desired level of quality and expected success. If the degree of self-efficacy becomes higher, people feel more confident in operation. In this case, if there are difficulties and obstacles, they are a motivating factor. Overcoming them reinforces desire to learn something new and increase one's skill. Conversely, people with low self-efficacy prefer to avoid difficult or potentially dangerous situations. Lack of self-efficacy increases stress and the risk of falling into a depression.

According to Bandura (1997) there are 4 sources of getting information about expected self-efficacy. They are:

1) Performance accomplishments;
2) Vicarious experience;
3) Verbal persuasion;
4) Physiological states.



Picture 6 Diagrammatic representation of the difference between efficacy expectations and outcome expectations.

Performance accomplishments are one of the most influential sources due to its base on personal mastery experiences. Individual successes increase mastery expectations while constant failures, on the contrary, decrease it. This effect especially occurs on the early stages of some course of events. A person with little experience in any activity, might be strongly influenced by failures. A series of failings will demotivate one to raise his/her mastery. Eventually a person will decide to abandon attempts and leave. On the contrary, experienced person is likely to have strong efficacy expectations. Regular practice and successes provide confidence which reduce the negative effect of the failure. Indeed, on the later stages of the practice some failures would seem to be even motivational, increasing passion and excitement of some activity. So, in general, the effects of failure on personal efficacy depend on time of practicing and experience in which the failure took place.

The previous source is not the only one for an individual to get information about his or her level of self-efficacy. Lots of expectations come from vicarious experience. While observing other people making threatening actions (treated dangerous by the individual but possibly not really hazardous) and not getting adverse consequences the individual can generate new expectations. “If they can do it, I’m also able to do the same”, an individual would think and try to persist in his activity. Vicarious experience is not so efficient as personal experience of success. It’s based on the social comparison so increase in efficacy is expected to be lower and more vulnerable to change.

Verbal persuasion is one of the most common and easily accessible ways of changing others behavior and expectations of self-efficacy. However simple telling to people of what to except proved to be weak form of influence. When people bear the long history of failures and dealing with them simple verbal motivation will be annihilated by a large set of negative experiences. “Numerous experiments have been conducted in which phobics receive desensitization treatment without any expectancy information or with suggestions that it is either highly efficacious or ineffective. The differential outcome expectations are verbally induced prior to, during, or immediately after treatment in the various studies. The findings generally show that desensitization reduces phobic behavior, but the outcome expectancy manipulations have either no effect or weak, inconsistent ones" (Bandura, 1997).

The last but not the least emotional arousal is another source of expectation change and self-efficacy analysis. While going through stressful of threatening situation an individual might summon emotional arousal which may or may not give informative value concerning personal competency. People, to some extent, rely on their psychological conditions when judging their anxiety and vulnerability to stress. High arousal, stress and nervous state of mind is more likely to diminish success in some activity. Consequently, people expect more success when they are calm, confident in themselves and the activities in which they are engaged. In case of success, this is confirmed by the practice that refers to the source of performance accomplishments, which further strengthens the individual's confidence (some synergy achieved) and increases its expectations and self-efficacy. Conversely, the presence of fear, nervousness, individuals tend to reflect on past failures and fears, which lowers the chances of success in ongoing activities and outputs of stress on a new, higher level than it was before. This might lead to the new failure and new securing of the negative template in mind of the individual.

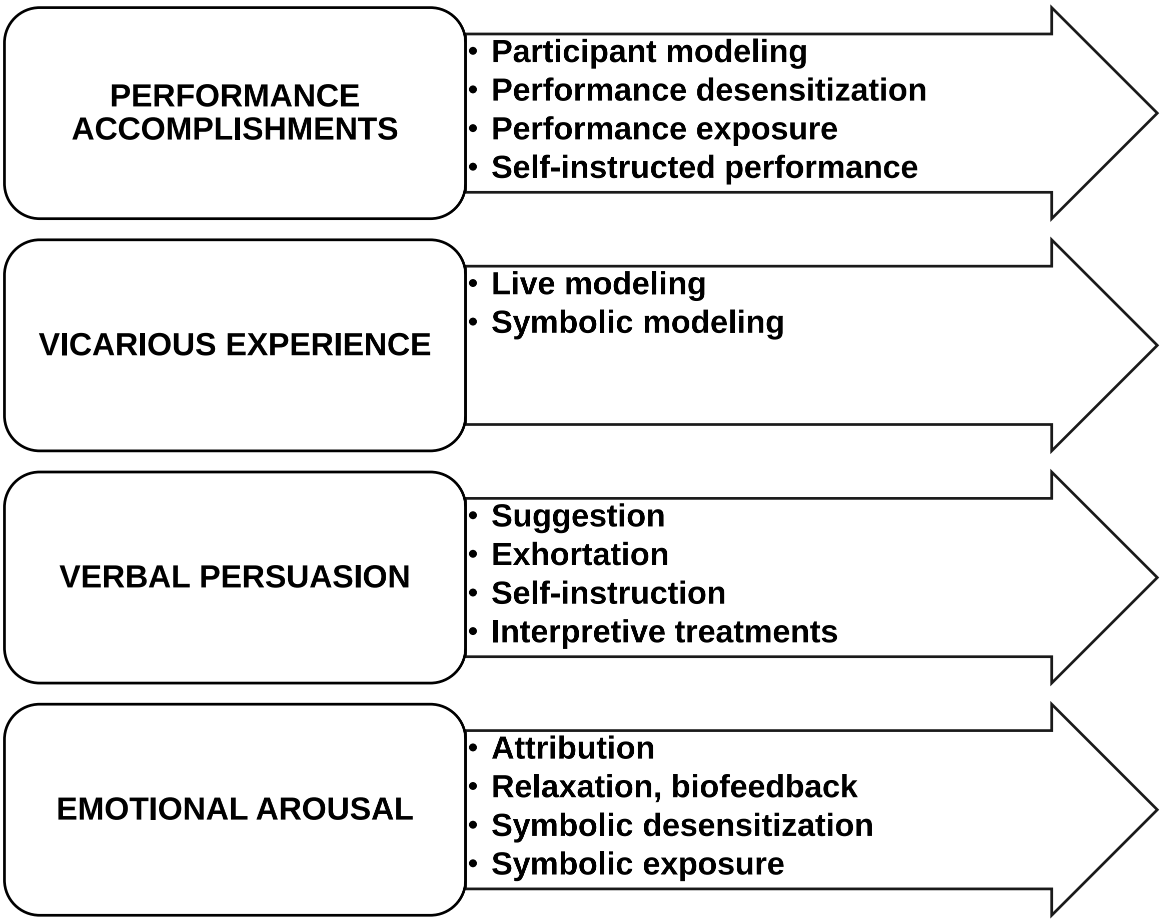
There are a number of factors that nullify or reduce the effect of successful experience. If the experience is contrary to long-established habit, its effect is much weaker.

One of these is the discrimination process. The individual may behave quite boldly, when he knows that the circumstances are safe, but in the real environment, the same procedure will be difficult or not performed at all due to fear. "People can gain competence through authentic means but, because of faulty appraisals of the circumstances under which they improve, will credit their achievements to external factors rather than to their own capabilities". Often people need to perform any action on their own, to gain self-confidence. Help-party model in this case will have a hindering effect.

Armed with what was covered above, we can investigate how different factors affect the changes in the market value of securities. One of these factors are rumors about possible events that appear from time to time in the market among traders. Although they are not reliable predictor, they can exert its influence, in varying degrees.

DiFonzo and Bordia (2002) completed two studies devoted to rumors effect on trading strategy. The first study discovered that in spite rumors had no association with prices, they made traders believe they were. Study 1 also found that rumors, and news, produced anti-regressive trading behavior as compared with controls, which was supported by the second study. In general, these studies show connections between rumor containing stable-cause explanations, causal attribution, anti-regressive prediction behavior, spurious association.

Talking about EMH researches often cover two points: first is the fact, that in the long run investors who behave irrationally will most probably lose their money and leave the market without having big influence on the prices. Second states that informational efficiency of securities markets guides firms and investors to efficient allocations of capital and labor. Irrational behavior of investors may have an impact on prices even with the existence of the EMH. Hirshleifer, Subrahmanyam, and Titman (2006) in their paper study this issues. "We show that when feedback from stock prices to cash flows is sufficiently strong, irrational investors can realize positive expected profits that exceed the expected profits of investors with fundamental information". Previous studies revealed that irrational traders are able to get higher profits by either bearing higher risks or exploiting private information more aggressively. In contrast to these arguments investors in authors' model earn positive expected profits without any private information that is inherently related to fundamentals, in a setting where risk-neutral market makers ensure that there is no market compensation for bearing risk. Further, these expected profits are inadvertently earned, in that they obtain in a setting where the irrational investors are price takers who naively ignore the feedback effect.



Picture 7 Major sources of efficacy information and the principal sources through which different modes of treatment operate

2.3. Technical analysis approach

As the research of Taylor and Allen (1992) showed at least 90% of investors treated technical analysis as useful analytical tool and used it to increase chances of creating profitable trading strategy. There also seemed to be, a clear consensus among respondents that chart analysis is used mostly as a guide to shorter-term exchange rate behavior and, moreover, that chartist advice should be used in conjunction with fundamentalist advice. Technical analysts, dealing with the minutiae of market changes, probably can get a good intuitive feel for a closer local approximation to the underlying economic structure. By doing so they gain popularity with traders whilst having no deep understanding of market forces-in the same way that a good billiards player may have no knowledge of physics.

Blume, Easley, and O'hara (1994) have investigated how technical analysis can bring benefits to traders in an economy in which the only uncertainty arises from the underlying information structure. In the model which they developed technical analysis was valuable because market statistics may be sufficient to reveal some information about and asset, but not all information. Because the underlying uncertainty in the economy was not resolved in one period, sequences of market statistics could provide information that was not impounded in a single market price. The most interesting results were obtained in delineating the important role played by volume. Volume provided information in a way distinct from that provided by price. Price impounded information about the average level of trader's private information. However unique to their model is the feature that volume captured the important information contained in the quality of traders’ information signals. Because the volume statistic was not normally distributed, if traders condition on volume they can sort out the information implicit in volume from that implicit in price. Authors have shown that volume plays a role beyond simply being a descriptive parameter of the trading process.

Their research focused on the quality, or precision, of information and suggested that the value of particular market statistics may vary depending upon characteristics of the information structure. Though the discussion was devoted to the potential applications of technical analysis for small, thinly followed stocks, it seems likely that even in active markets volume could have played an important role.

Another article (Neftci, 1991) discussed some criteria that one can apply in evaluating the set of ad hoc prediction rules widely used in financial markets and generally referred to as technical analysis. There have been shown that a few of these rules generate well-defined techniques of forecasting. Under the hypothesis, economic time series are Gaussian, and even well-defined rules were shown to be useless in prediction. At the same time, the discussion indicated that if the processes under consideration were nonlinear, then the rules of technical analysis might capture some information ignored by Wiener-Kolmogorov prediction theory. Tests done using the Dow-Jones industrials for 1911-76 suggested that this may indeed be the case for the moving average rule.

Lo and MacKinlay (1988) found statistically significant evidence that the price movements are not random. Exploring the weekly data, they compared the logarithms of the variances for periods of one and four weeks in the period from 1962 to 1985. According to the results of tests they rejected the hypothesis of a random walk, showing the presence of positive autocorrelation.

2.4. Fundamental analysis predictions

Dividend yield as an explaining variable was taken into consideration by Fama and French (1988). They have taken dividend/price ratios (D/P) or dividend yield to predict returns of NYSE portfolios which were equally weighted. Return horizons (holding periods) varied from one month to four years. Their regression of returns on dividend yields revealed that "time variation in expected returns accounts for small fractions of the variances of short-horizon returns". The shorter return horizon was analyzed, the smaller was the explained variation. For example, when they took monthly and quarterly returns dividend yields typically explained less than 5% of the variances. However, dividend yields often predicted more than 25% of the variances of two- to four-year returns. Authors gave the following explanation of this fact: "The persistence (high positive autocorrelation) of expected returns causes the variance of expected returns, measured by the fitted value in the regressions of returns on dividend yields, to grow more than in proportion to the return horizon. On the other hand, the growth of the variance of the regression residuals is attenuated by a discount-rate effect: shocks to expected returns are associated with opposite shocks to current prices. The cumulative price effect of an expected return shock and the associated price shock is roughly zero. On average, the expected future price increases implied by higher expected returns are just offset by the immediate decline in the current price. Thus the time variation of expected returns gives rise to mean-reverting or temporary components of prices".

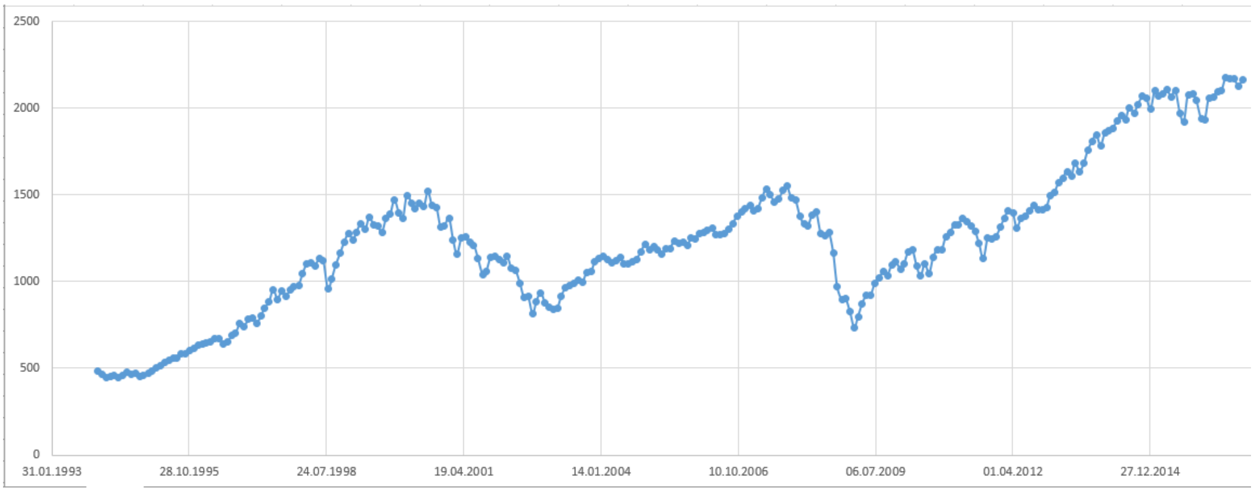
Price to earnings analysis is and another attempt to predict stock returns' fluctuations. Campbell and Shiller (1988) in their article used history data of accounting earnings to evaluate present value of future dividends. And they succeeded showing that "a long moving average of real earnings helps to forecast future real dividends. The ratio of this earnings variable to the current stock price is a powerful predictor of the return on stock, particularly when the return is measured over several years". However, Malkiel (2003) doubts in the correctness of this kind of approach. He proposes to consider "the recent experience of investors who have attempted to undertake investment strategies based either on the level of the price-earnings multiple or the dividend yield to predict future long horizon returns. Price-earnings multiples for the Standard & Poor's 500 stock index rose into the low 20s on June 30, 1987 (suggesting very low long horizon returns). Dividend yields fell below three percent. The average annual total return from the index over the next 10 years was an extraordinarily generous 16,7 percent. Dividend yields, again, fell to three percent in June of 1992. Price-earnings multiples rose to the mid-twenties. The subsequent return through June 2002 was 11,4 percent. The yield of the index fluctuated between two and three percent from 1993 through 1995 and earnings multiples remained in the mid-twenties, yet long horizon returns through June 30, 2002 fluctuated between 11 and 12 percent. Even from early December 1996, the date of Campbell and Shiller's presentation to the Federal Reserve suggesting near zero returns for the S&P500, the index provided almost a seven percent annual return through mid-2002". According to the author such results should be treated with the great caution and should be double checked if one is going to use them in order to predict market returns.

3. Methodology

In order to test the hypothesis about randomness of stock prices movement the series of test will be conducted in Master thesis. Will be analyzed behavior of several stock indexes. In order to do so I'll undertaken steps which had been used by other researches, e.g. Poshakwale (1996), Lo and MacKinlay (1988) and others.

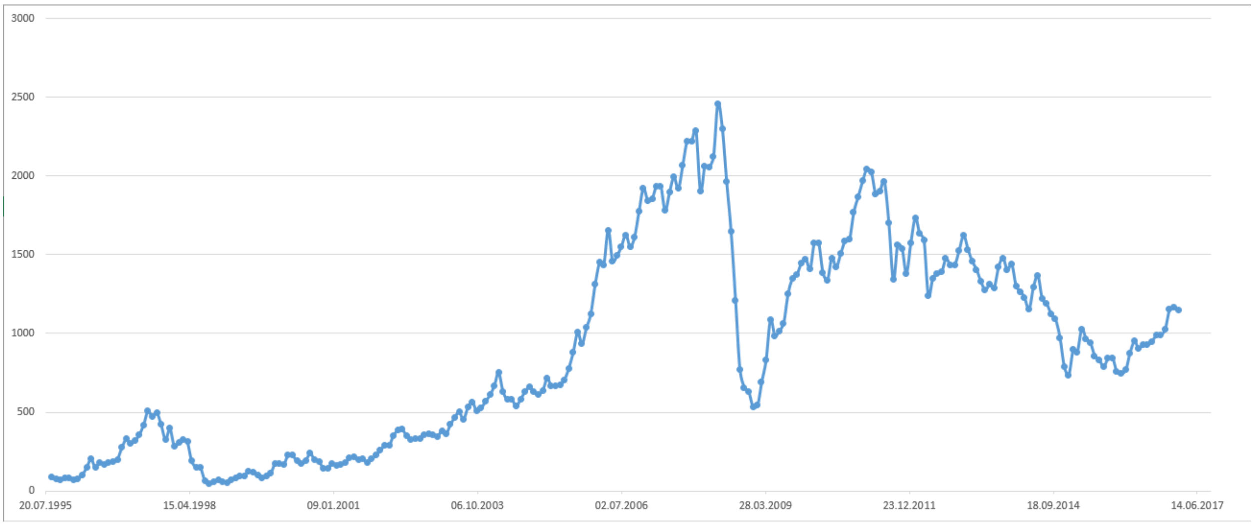
I've obtained historical prices of several indexes to test the hypothesis about the randomness of stock price movement. They are Standard & Poor's 500 (S&P500), Oslo Børs (OSEAX) and Russian Trading System (RTS). I've taken the following timing periods:

- Monthly data for the period 01.1994 to 12.2016 for S&P500.



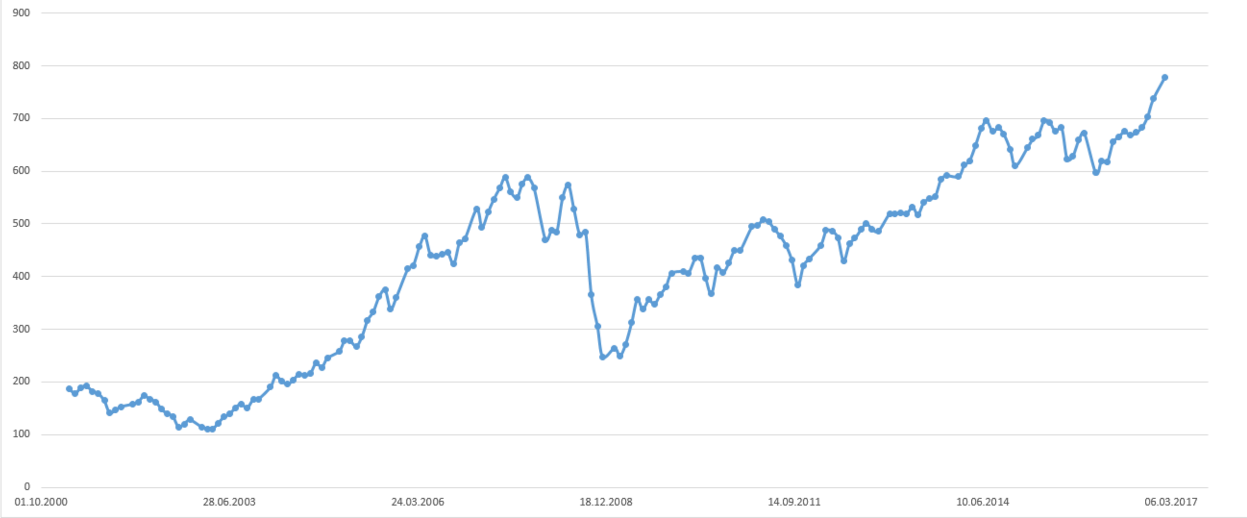
Picture 8 S&P500 prices index

- Monthly data for the period 09.1995 to 02.2017 for RTS.



Picture 9 RTS prices index

- Monthly data for the period 05.2002 to 12.2016 for OSEAX.



Picture 10 OSEAX prices index

Source of information is YAHOO Finance service.

It is necessary to conduct those test which would provide best reflection of characteristics of the time series. They also help to assess randomness of stock price changes. I've chosen the following ones:

- Jarque-Bera Normality Test;
- Augmented Dickey-Fuller test;
- Kolmogorov Smirnov Goodness of Fit Test;
- Serial Correlation Coefficients Test and some more.

One of the basic assumptions of efficient market hypothesis is the normality of the returns distribution (Poshakwale, 1996). In order to determine whether the distribution is normal, it is necessary to evaluate its characteristics of skewness and kurtosis.

Kolmogorov-Smirnov test is used to test if the sample follows some distribution law, to check that the empirical distribution consistent with the proposed model. In our case I suggested comparing with standard normal distribution. The test was based on comparison of the sample's cumulative distribution against the standard cumulative distribution.

Serial Correlation Test is one of the basic tests for efficient market hypothesis checking. It provides a reliable estimation of whether the variables in the series are dependent or independent. In order to perform the test, I transformed the series by taking the first difference and computing the autocorrelation.

Another interesting approach was carried out by Lo and MacKinlay (1988): "The plausibility of the random walk model may be checked by comparing the variance estimate of X_t

- X_{t-1} to, say, one-half the variance estimate of X_t - X_{t-2}. This is the essence of specification test". The null hypothesis H0 was about disturbance term ε and its iid properties.

$$H: \varepsilon_t \; i.i.d. N(0, \sigma_0^2)$$

Then several estimators for unknown parameters μ and σ were introduced:

$$\hat{\mu} \equiv \frac{1}{2n}(X_{2n} - X_0)$$

$$\hat{\sigma}_a^2 \equiv \frac{1}{2n}\sum_{k-1}^{2n}(X_k - X_{k-1} - \hat{\mu})^2$$

$$\hat{\sigma}_b^2 \equiv \frac{1}{2n}\sum_{k-1}^{2n}(X_{2k} - X_{2k-1} - 2\hat{\mu})^2$$

After which other estimators were defined:

$$J_d(q) \equiv \hat{\sigma}_b^2(q) - \hat{\sigma}_a^2$$

$$J_r(q) \equiv \frac{\hat{\sigma}_b^2(q)}{\hat{\sigma}_a^2} - 1$$

That have been used for hypothesis testing

One remarkable observation about these investigations is that both of them rejected the null hypothesis about random walk in stock prices in some extent. However, both articles claim that it's not the proof of market inefficiency. It just "imposes restrictions upon the set of plausible economic models for asset pricing".

Another test is the augmented Dickey-Fuller (ADF) test which I use to check the existence of a unit root in the series of price movements in the stock index series. I used the following equation through OLS:

$$\Delta P_t = \alpha_0 + \alpha_1 t + \rho_0 P_{t-1} + \sum_{i=1}^{q} \rho_i \Delta P_{it-i} + \varepsilon_{it}$$

Where P_t is price at moment t, $\Delta P_t = P_t - P_{t-1}$, ρ_i are coefficients to be estimated. q is the number of lagged terms, t is the trend term, α_i is the estimated coefficient for the trend, α_0 is the constant and finally ε is white noise. The H_0 of a random walk implies that $H_0: \rho_0 = 0$. The alternative hypothesis claims that $H_1: \rho_0 \neq 0$. If I fail to reject H_0 this means that I cannot reject that time series has the properties of random walk (Borges, 2011).

After conducting series of tests for each set of data and receiving the results, it will be possible to make a comparison of their features, see the repetitive and different patterns of their behavior. Based on the results it can be concluded that the hypothesis of the random walk is valid for the selected indexes.

4. Data analysis

For checking the independence in stock returns I used Runs test. It determines whether successive price changes are dependent or independent of each other. Under the null hypothesis of random walk they should be independent. I test the null hypothesis through observing the number of runs of price changes with the same signs. I consider two approaches: in the first, I define as a positive return (+) any return greater than zero, and a negative return (-) if it is below zero; in the second approach, we classify each return according to its position with respect to the mean return of the period under analysis. In this last approach, I have a positive (+) each time the return is above the mean return and a negative (-) if it is below the mean return (Borges, 2011). This second approach has the advantage of allowing for and correcting the effect of an eventual time drift in the series of returns. Worth noting that this is a non-parametric test. It does not require the returns to be normally distributed. The runs test is based on the premise that if price changes (returns) are random, the actual number of runs (R) should be close to the expected number of runs (μ_R).

I mark number of positive runs with n_+ and number of negative runs with n_-. Totsl number of observations is equal $n = n_+ + n_-$. For large sample sizes, the test statistic is approximately normally distributed:

$$Z = \frac{R - \mu_R}{\sigma_R} \approx N(0,1)$$

Where

$$\mu_R = \frac{2n_+n_-}{n} + 1 \text{ and } \sigma_R = \sqrt{\frac{2n_+n_-(2n_+n_- - n)}{n^2(n-1)}}$$

In order to test if the observed distribution fit theoretical normal or uniform distribution I will use non-parametric test: Kolmogorov Smirnov Goodness of Fitness Test (KS). It is used to determine how good a random sample of data fits some kind of distribution (e.g. uniform, normal or Poisson). The test is based on comparison of the sample’s cumulative distribution against the standard cumulative function for each distribution. The Kolmogorov-Smirnov one sample goodness of fit test compares the cumulative distribution function for a variable with a uniform or normal distributions and tests whether the distributions are homogeneous. I use both normal and uniform parameters to test distribution (Poshakwale, 1996).

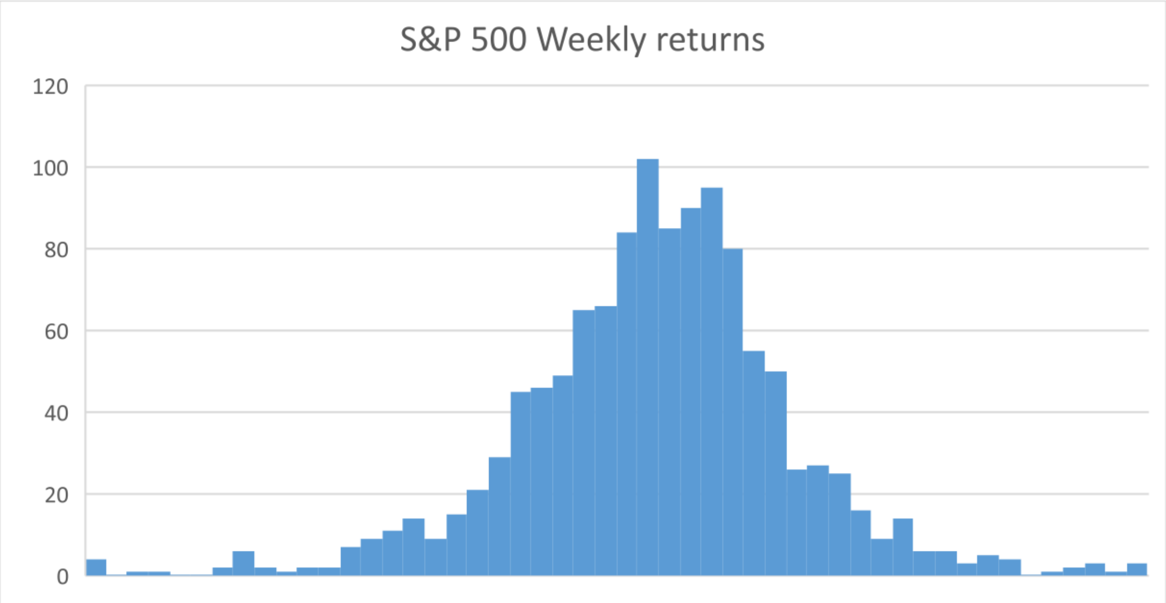
In order to test the Efficient Market Hypothesis (EMH) in the weak form, Serial Correlation Coefficient Test is frequently used. The Serial Correlation Coefficient studies the relationship between the values of a random variable at particular point of time and its value in the previous period. The population serial correlation (P_a) coefficient is estimated using the sample serial correlation coefficient (R_a). For complete independence $P_a = 0$, a significant test may be performed on the variation of R_a from 0. Here confidence intervals of two and three standard errors are used.

Autocorrelations are reliable measures for testing of dependence/independence of random variables in a series. If no autocorrelations are found in a series then the series is considered random. We transform the series by taking the first difference and compute the autocorrelations. The autocorrelation coefficients have been computed for the transformed index in order to establish whether information is obtained even with transformation of the higher order.

4.1. Descriptive statistics

One of the major assumptions that implies the random walk theory and, therefore, EMH is that the distribution of stock prices should be normal in order to be random. Any normal distribution is an advantage because I'll only two summary measures, mean and variance, to describe the entire distribution. The normality of distribution is also one of the basic assumptions underlying the capital asset pricing models (Poshakwale, 1996).

I've constructed the histograms of three index and compared them with the normal distribution curve (see below). In case the distribution has more cases but not symmetric or if one of the "tails" is longer than the other is called "skewed". Positively skewed distribution has longer right tail and moved to the larger values, and vice versa, negative skew means longer left tail. Kurtosis indicates the extent to which, for a given standard deviation, observations cluster around a central point. If cases within a distribution cluster more than those in the normal distribution (that is the distribution is more peaked), the distribution is called leptokurtic. If cases cluster less than in the normal distribution (that is, it is flatter), the distribution is termed platokurtic.

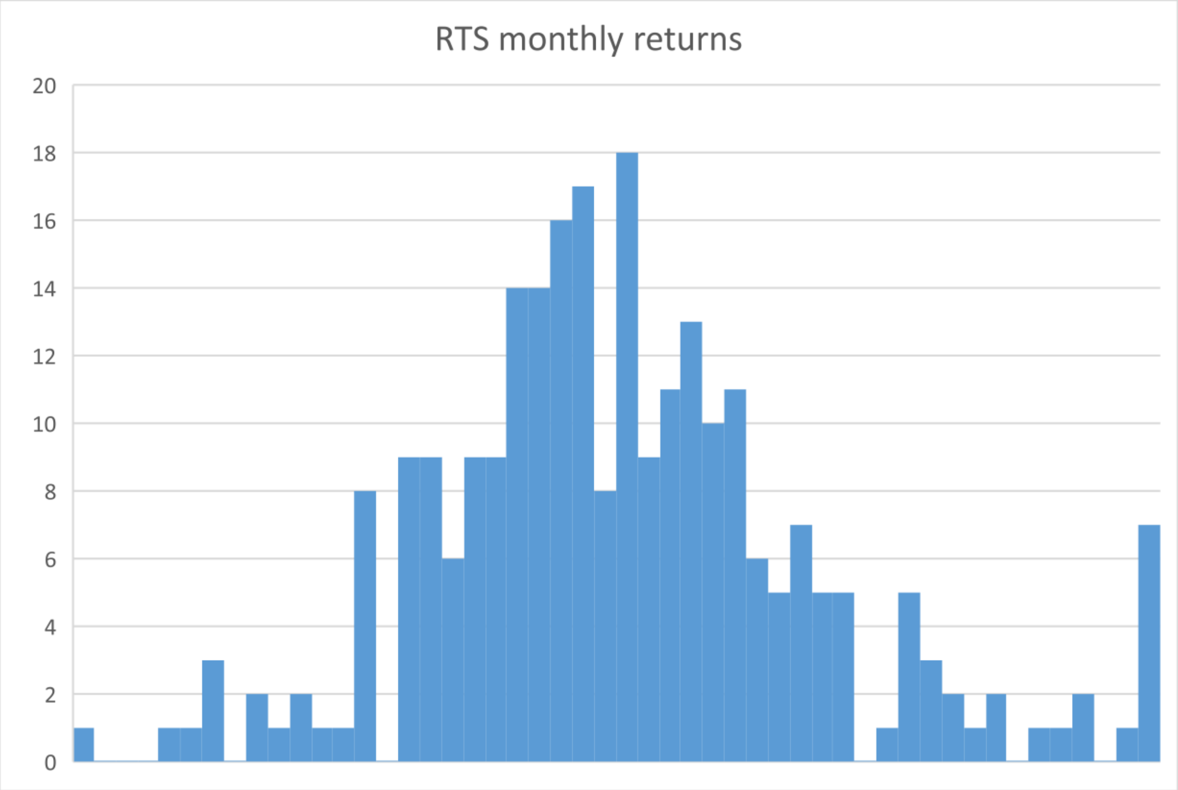


Histogram 1 S&P 500 Weekly returns (01.1994 – 12.2016)

Skewness	- 0,49
Kurtosis	5,31
Number of obs.	1199
St. dev	0,02

Mean	0,002

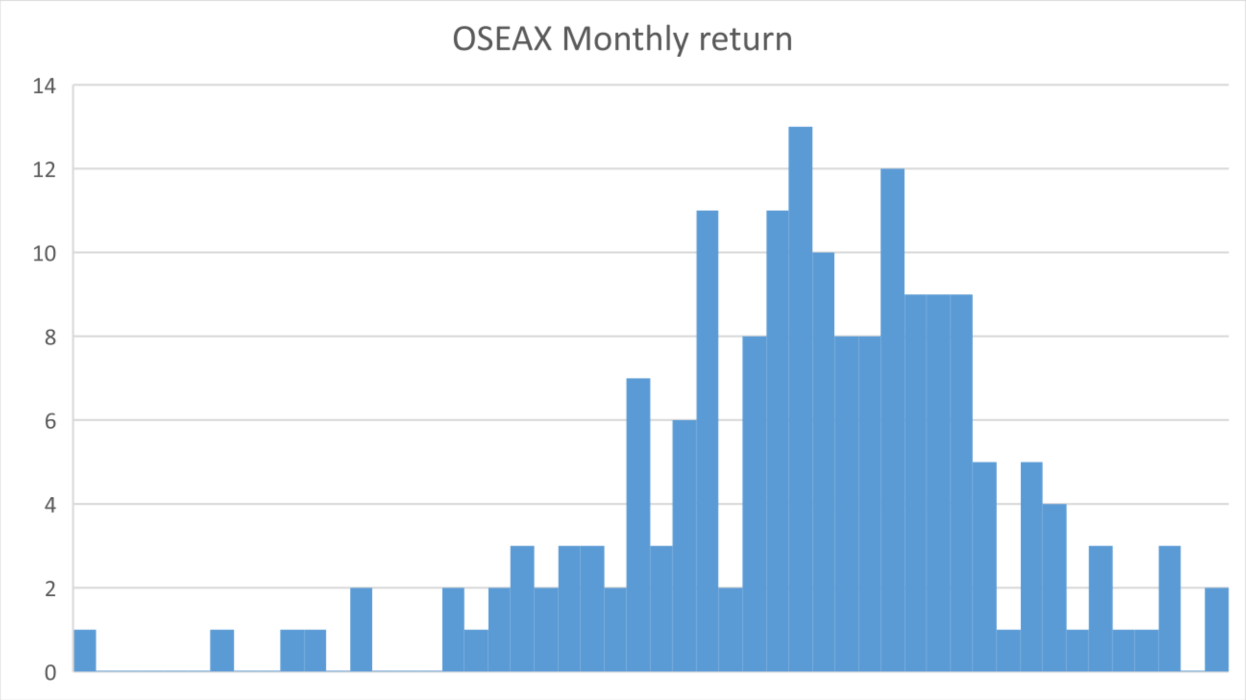
On the histogram above we can see that the distribution is not normal, it's negatively skewed and has very large kurtosis. However, I cannot reject the random walk hypothesis using only the descriptive statistics, further investigation is needed and will be presented below. Now let's see the other indexes.



Histogram 2 RTS Monthly returns (09.1995 – 02.2017)

Skewness	2,84
Kurtosis	18,89
Number of obs.	258,00
St. dev	0,16
Mean	0,0002

The same or even worse situation here on monthly returns of RTS. Very strong kurtosis accompanied by positive skewness leads us to rejection of random walk. However, I still need to perform some more tests.



Histogram 3 OSEAX Monthly returns (03.2001 – 02.2017)

Skewness	- 0,69
Kurtosis	1,36
Number of obs.	176,00
St. dev	0,07
Mean	0,0103

Finally, OSEAX returns do not allow us to suspect the normal distribution here either. Negative skewness and kurtosis smaller than 3 makes the distribution looks differently than the standard normal distribution does.

I begin my analysis from S&P500 index. The data was downloaded from YAHOO Finance and prepared for the analysis using resources and tools of RStudio and Excel. The code and results are presented below:

4.2. S&P 500 testing

```
####################### S&P 500 testing ###############################

View(SP500)

sp_ret = ts(log(Close[1:275])/log(Close[2:276]), start=c(1994,1), end =c(2015,5), frequency = 12)
plot(Close)
plot(sp_ret)
hist(sp_ret, breaks = 100, freq=TRUE)
sp_ret

attach(SP500)

#### Ljung-Box test ####
Box.test(sp_ret, lag = 1, type = "Ljung") #H0 rejected
```

```
#### Jarque-Bera normality test ####
jarque.bera.test(sp_ret)        #H0 about normality is rejected

#### Dickey-Fuller test for stability of a time series variable ####
library(urca) #Get correlogram check lag order
adf.sp = ur.df(sp_ret, type = c("none"), lags=1)
summary(adf.sp)   #Ho about stability is rejected
plot(adf.sp)

#### Kolmogorov-Smirnov Tests ###
set.seed(3000)
xseq<-seq(-4,4,.01)
ks.test(sp_ret,pnorm(xseq, 0, 1))   #Reject the H0 that SP returns follow standard
normal distribution
```

I developed the series of tests in order to investigate the characteristics of this time series. In particular, we are interested whether there are autocorrelation coefficients that are jointly significantly different from zero, whether the characteristics of skewness and kurtosis are similar to the normal, whether the time series is stable and so on.

After running this code, I obtained the following results:

```
Box-Ljung test

data:  sp_ret
X-squared = 11.639, df = 1, p-value = 0.0006457
```

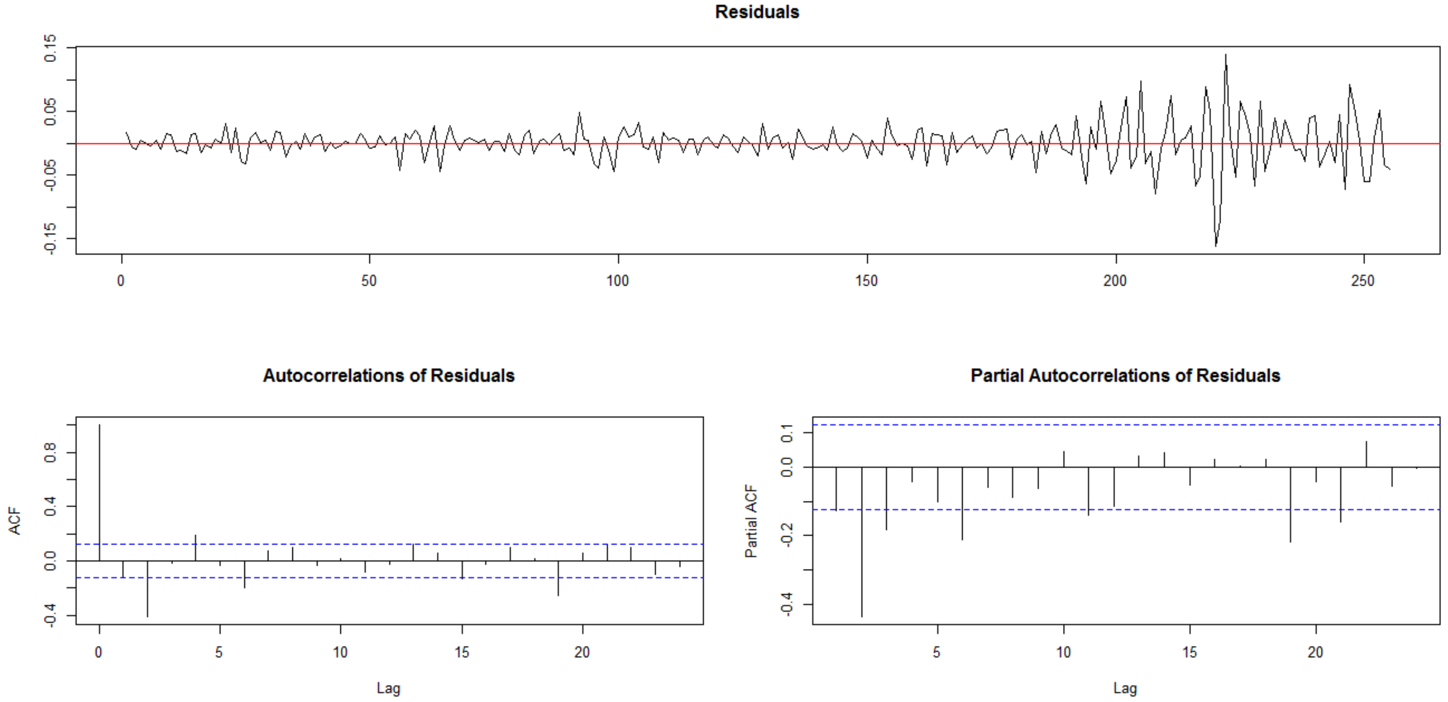
Conclusion: we reject the null hypothesis about no autocorrelation.

```
Jarque Bera Test

data:  sp_ret
X-squared = 756, df = 2, p-value < 2.2e-16
```

Conclusion: we reject the null hypothesis about normality.

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.162072 -0.011547  0.002366  0.013595  0.140745 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
z.lag.1    -0.0005024  0.0019175  -0.262    0.794    
z.diff.lag -0.3239746  0.0595507  -5.440 1.25e-07 ***
---
```

```
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0307 on 253 degrees of freedom
Multiple R-squared:  0.1051,       Adjusted R-squared:  0.09803
F-statistic: 14.86 on 2 and 253 DF,  p-value: 7.925e-07


Value of test-statistic is: -0.262

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```



Picture 11 ADF test for S&P500

Conclusion: we reject the null hypothesis about stability.

```
Two-sample Kolmogorov-Smirnov test

data:  sp_ret and pnorm(xseq, 0, 1)
D = 0.68437, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Conclusion: we reject the null hypothesis that SP returns follow standard normal distribution.

Runs test

The following Runs test was presented in Excel. Instead of using monthly data I used 1199 weekly observations for the same period. The results are presented below:

Symbol	Description	Value
Mean	Average return for the period	0,002
R	Number of runs in sample	639

n_-	Number of observations below the average	578
n_+	Number of observations above the average	621
n	Total number of observations	1199
E (R)	Expected number of runs	600
Var (R)	Variance	298,73
StDev (R)	Standard deviation	17,28
Z	Calculated Z-value	2,27
p-value	**Probability value**	**0,0302**

As we can see the probability is below 5-percent significance level so the random walk hypothesis can be rejected in this test.

4.3. RTS testing

Data analysis was continued with RTS index. The code is presented below:

```
View(RTS)
attach(RTS)

rts_ret = ts(log(Close[1:257])/log(Close[2:258]), start=c(1995,9), end =c(2017,2), frequency = 12)
plot(Close)
plot(rts_ret)
hist(rts_ret, breaks = 100, freq=TRUE)
rts_ret

#---- Ljung-Box test -----#
Box.test(rts_ret, lag = 1, type = "Ljung") #H0 rejected

#---- Jarque-Bera normality test  -----#
jarque.bera.test(rts_ret)       #H0 about normality is rejected

#---- Dickey-Fuller test for stability of a time series variable  -----#
library(urca) #Get correlogram check lag order
adf.sp = ur.df(rts_ret, type = c("none"), lags=1)
summary(adf.sp)   #Ho about stability is rejected
plot(adf.sp)

#---- Kolmogorov-Smirnov Tests  -----#
set.seed(3000)
xseq<-seq(-4,4,.01)
ks.test(rts_ret,pnorm(xseq, 0, 1))    #Reject the H0 that SP returns follow standard normal distribution
```

I developed the series of tests in order to investigate the characteristics of this time series. In particular, we are interested whether there are autocorrelation coefficients that are jointly significantly different from zero, whether the characteristics of skewness and kurtosis are similar to the normal, whether the time series is stable and so on.

After running this code, I obtained the following results:

```
Box-Ljung test
```

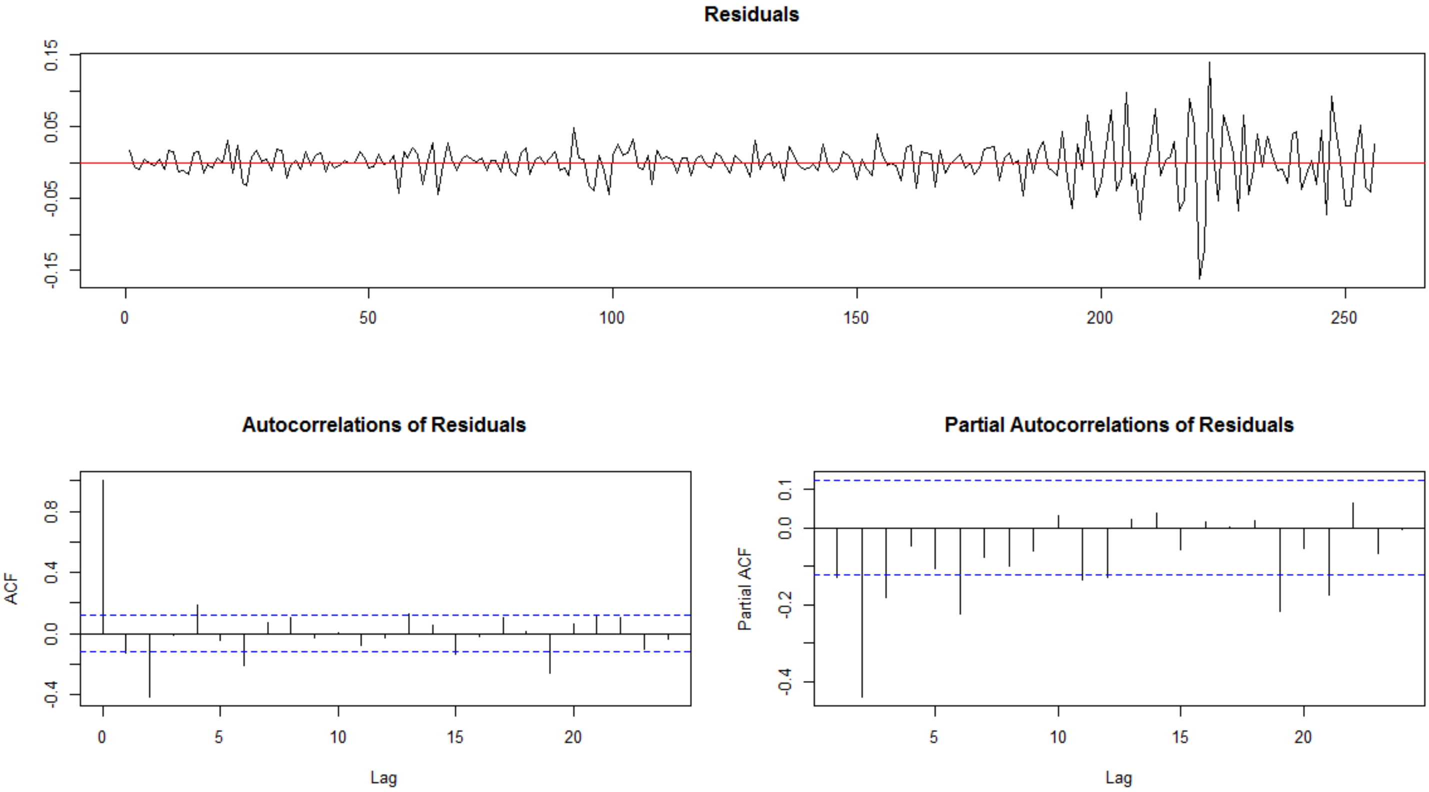
```
data:  rts_ret
X-squared = 11.788, df = 1, p-value = 0.0005963
```

Conclusion: we reject the null hypothesis about no autocorrelation.

```
Jarque Bera Test

data:  rts_ret
X-squared = 765.64, df = 2, p-value < 2.2e-16
```

Conclusion: we reject the null hypothesis about normality.

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.162129 -0.011609  0.002286  0.013673  0.140498 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
z.lag.1    -0.0004063  0.0019128  -0.212    0.832    
z.diff.lag -0.3264212  0.0594389  -5.492 9.65e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03068 on 254 degrees of freedom
Multiple R-squared:  0.1064,	Adjusted R-squared:  0.09934 
F-statistic: 15.12 on 2 and 254 DF,  p-value: 6.259e-07


Value of test-statistic is: -0.2124 

Critical values for test statistics: 
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```



Picture 12 ADF test for RTS

Conclusion: we reject the null hypothesis about stability.

```
Two-sample Kolmogorov-Smirnov test

data:  rts_ret and pnorm(xseq, 0, 1)
D = 0.68448, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Conclusion: we reject the null hypothesis that RTS returns follow standard normal distribution.

Runs test

Symbol	Description	Value
Mean	Average return for the period	-0,0002
R	Number of runs in sample	503,00
n_-	Number of observations below the average	591
n_+	Number of observations above the average	522
n	Total number of observations	1113
E (R)	Expected number of runs	555
Var (R)	Variance	275,87
StDev (R)	Standard deviation	16,61
Z	Calculated Z-value	- 3,15
p-value	**Probability value**	**0,0028**

In case of RTS index I can reject the random walk hypothesis on 5 and 1 percent level.

4.4. OSEAX testing

Data analysis was continued with RTS index. The code is presented below:

```
View(OSEAX)
attach(OSEAX)

os_ret = ts(log(Close[1:175])/log(Close[2:176]), start=c(2002,5), end =c(2016,12), frequency = 12)
plot(Close)
plot(os_ret)
hist(os_ret, breaks = 100, freq=TRUE)
os_ret

#---- Ljung-Box test -----#
Box.test(os_ret, lag = 1, type = "Ljung") #H0 rejected

#---- Jarque-Bera normality test -----#
jarque.bera.test(os_ret)        #H0 about normality is rejected

#---- Dickey-Fuller test for stability of a time series variable -----#
library(urca) #Get correlogram check lag order
adf.sp = ur.df(os_ret, type = c("none"), lags=1)
summary(adf.sp)   #Ho about stability is rejected
plot(adf.sp)

#---- Kolmogorov-Smirnov Tests -----#
set.seed(3000)
xseq<-seq(-4,4,.01)
ks.test(os_ret,pnorm(xseq, 0, 1))  #Reject the H0 that SP returns follow standard normal distribution
```

I developed the series of tests in order to investigate the characteristics of this time series. In particular, we are interested whether there are autocorrelation coefficients that are jointly significantly different from zero, whether the characteristics of skewness and kurtosis are similar to the normal, whether the time series is stable and so on.

After running this code, I obtained the following results:

```
Box-Ljung test

data:  os_ret
X-squared = 6.5326, df = 1, p-value = 0.01059
```

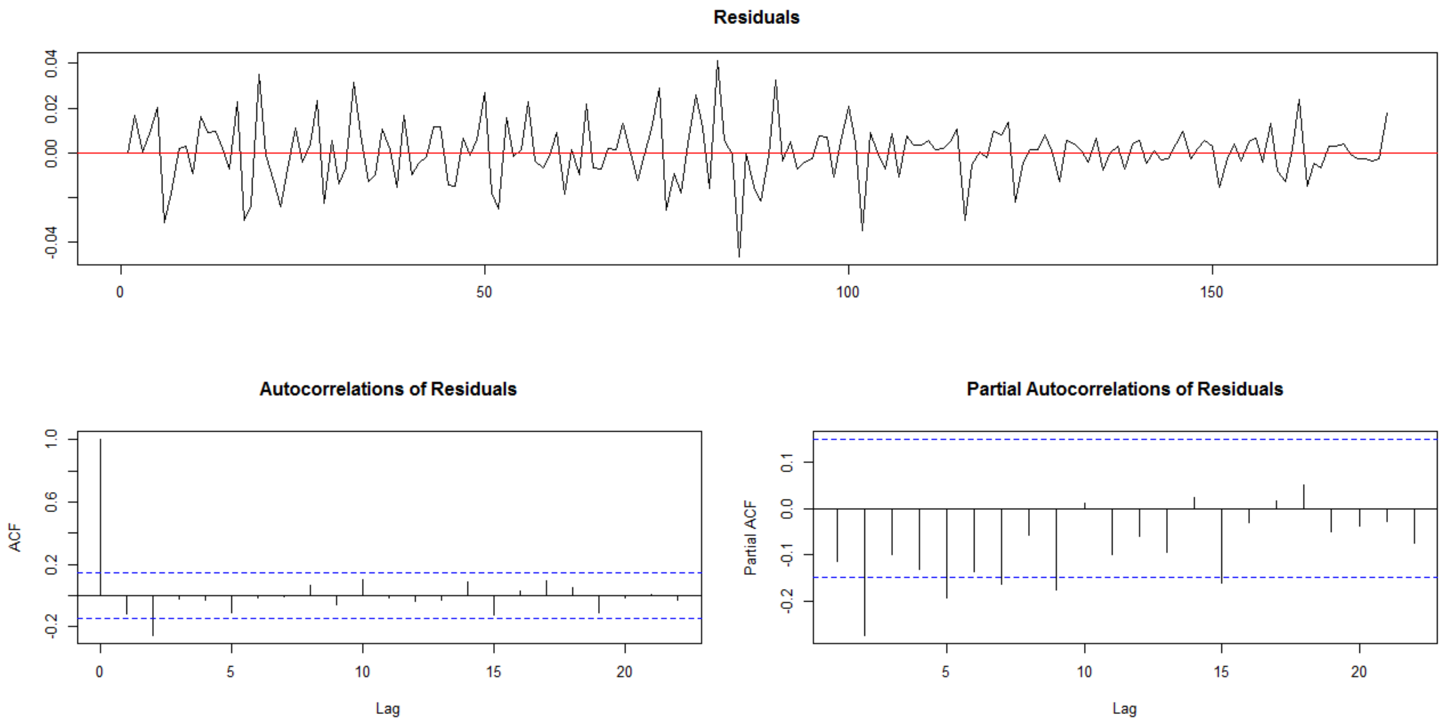
Conclusion: we reject the null hypothesis about no autocorrelation.

```
Jarque Bera Test

data:  os_ret
X-squared = 58.297, df = 2, p-value = 2.193e-13
```

Conclusion: we reject the null hypothesis about normality.

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################
```

```
Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.046615 -0.007060  0.000861  0.006395  0.041420 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
z.lag.1     1.835e-05  1.018e-03   0.018    0.986    
z.diff.lag -4.452e-01  6.811e-02  -6.537 6.85e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01341 on 172 degrees of freedom
Multiple R-squared:  0.199,	Adjusted R-squared:  0.1897 
F-statistic: 21.37 on 2 and 172 DF,  p-value: 5.142e-09


Value of test-statistic is: 0.018 

Critical values for test statistics: 
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```



Picture 9 ADF test for OSEAX

Conclusion: we reject the null hypothesis about stability.

`Two-sample Kolmogorov-Smirnov test`

```
data:  os_ret and pnorm(xseq, 0, 1)
D = 0.74407, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Conclusion: we reject the null hypothesis that OSEAX returns follow standard normal distribution.

Runs test

Symbol	Description	Value
Mean	Average return for the period	0,0103
R	Number of runs in sample	76,00
n_-	Number of observations below the average	82
n_+	Number of observations above the average	94
n	Total number of observations	176
E (R)	Expected number of runs	89
Var (R)	Variance	43,34
StDev (R)	Standard deviation	6,58
Z	Calculated Z-value	- 1,91
p-value	**Probability value**	**0,0641**

Here I failed to reject the hypothesis even on 5 percent level.

4.5. Lo & MacKinlay specification test

In this section I will try to reperform test that was conducted by Lo and MacKinlay (1988). I've obtained both weekly and monthly observations of S&P500 (1199 and 276 observations respectively). I consider the following estimators for the unknown parameters:

$$\hat{\mu} \equiv \frac{1}{2n}(X_{2n} - X_0)$$

$$\hat{\sigma}_a^2 \equiv \frac{1}{2n}\sum_{k-1}^{2n}(X_k - X_{k-1} - \hat{\mu})^2$$

$$\hat{\sigma}_b^2 \equiv \frac{1}{2n}\sum_{k-1}^{2n}(X_{2k} - X_{2k-1} - 2\hat{\mu})^2$$

The estimators $\hat{\mu}$ and $\hat{\sigma}_a^2$ correspond to the maximum-likelihood estimators of the $\hat{\mu}$ and $\hat{\sigma}_a^2$ parameters: $\hat{\sigma}_b^2$ is also an estimator of $\hat{\sigma}_0^2$ but uses only the subset of n + 1 observations X_0, $X_2, X_4 \dots X_{2n}$ and corresponds formally to 1/2 times the variance estimator for increments of even-numbered observations. According to standard asymptotic theory, all three estimators are strongly consistent; In other words, this means that holding all other parameters constant, as the total number of observations 2n increases without bound the estimators converge almost surely to their population values. In addition, it is well known that both $\hat{\sigma}_a^2$ and $\hat{\sigma}_b^2$ possess the following Gaussian limiting distributions:

$$\sqrt{2n}(\hat{\sigma}_a^2 - \hat{\sigma}_0^2) \sim N(0, 2\sigma_0^4)$$

$$\sqrt{2n}(\hat{\sigma}_b^2 - \hat{\sigma}_0^2) \sim N(0, 4\sigma_0^4)$$

In this case distributional equivalence is asymptotic. It can be easily demonstrated that differences of the variances are asymptotically Gaussian with zero mean. However the variance of the limiting distribution is not clear cause those two variance estimators are not asymptotically uncorrelated. But since the estimator $\hat{\sigma}_a^2$ is asymptotically efficient under the null hypothesis H, we may apply Hausman (1978) result, which shows that the asymptotic variance of the difference is simply the difference of the asymptotic variances. We can define $J_d \equiv \hat{\sigma}_b^2 - \hat{\sigma}_a^2$ and then we have the result

$$\sqrt{2n}J_d \sim N(0, 2\sigma_0^4)$$

Using any consistent estimator of the asymptotic variance of J_d, a standard significance test may be performed. I will use slightly another way using the test statistic provided by ratio of the variances J_r:

$$J_r \equiv \frac{\hat{\sigma}_b^2}{\hat{\sigma}_a^2} - 1$$

$$\sqrt{2n}J_r \sim N(0, 2)$$

Although the variance estimator $\hat{\sigma}_b^2$ is based on the differences of every other observation, alternative variance estimators may be obtained by using the differences of every q_{th} observation. For instance, we obtain nq + 1 observations X_0, $X_1, X_2 \ldots X_{nq}$ where q is any integer greater than 1. Define the estimators:

$$\hat{\mu} \equiv \frac{1}{nq}\sum_{k-1}^{nq}(X_k - X_{k-1}) = \frac{1}{nq}(X_{nq} - X_0)$$

$$\hat{\sigma}_a^2 \equiv \frac{1}{nq}\sum_{k-1}^{nq}(X_k - X_{k-1} - \hat{\mu})^2$$

$$\hat{\sigma}_b^2(q) \equiv \frac{1}{nq}\sum_{k-1}^{n}\left(X_{qk} - X_{qk-q} - q\hat{\mu}\right)^2$$

$$J_d(q) \equiv \hat{\sigma}_b^2(q) - \hat{\sigma}_a^2$$

$$J_r(q) \equiv \frac{\hat{\sigma}_b^2(q)}{\hat{\sigma}_a^2} - 1$$

The first specification test made by authors was performed using Theorem 1:

Theorem 1. Under the null hypothesis H, the asymptotic distributions of $J_d(q)$ and $J_r(q)$ are given by

$$\sqrt{nq}J_d(q) \sim N(0, 2(q-1)\sigma_0^4)$$

$$\sqrt{nq}J_r(q) \sim N(0, 2(q-1))$$

The next pair of refinements of the statistics J_d and J_r result in more good looking finite-sample properties. The first is to use overlapping q_{th} differences of X_t, in estimating the variances by defining the following estimator of σ_0^2:

$$\hat{\sigma}_c^2(q) \equiv \frac{1}{nq^2}\sum_{k=q}^{nq}\left(X_k - X_{k-q} - q\hat{\mu}\right)^2$$

Previous expression differs from estimator $\hat{\sigma}_b^2(q)$ due to this sum has nq-q+1 terms, whereas the estimator $\hat{\sigma}_b^2(q)$ contains only n terms. By using overlapping q_{th} increments, we obtain a more efficient estimator and hence a more powerful test. Using $\hat{\sigma}_c^2(q)$ in our variance-ratio test, we define the corresponding test statistics for the difference and the ratio as

$$M_d(q) \equiv \hat{\sigma}_c^2(q) - \hat{\sigma}_a^2$$

$$M_r(q) \equiv \frac{\hat{\sigma}_c^2(q)}{\hat{\sigma}_a^2} - 1$$

This last refinement includes using unbiased variance estimators in our calculation of the M-statistics. Denote the unbiased estimators as $\bar{\sigma}_a^2$ and $\bar{\sigma}_c^2$ where

$$\bar{\sigma}_a^2 = \frac{1}{nq-1}\sum_{k=1}^{nq}(X_k - X_{k-1} - \hat{\mu})^2$$

$$\bar{\sigma}_c^2(q) = \frac{1}{m}\sum_{k=1}^{nq}\left(X_k - X_{k-q} - \widehat{q\mu}\right)^2$$

$$m = q(nq - q + q)\left(1 - \frac{q}{nq}\right)$$

Now I'm able to define the test-statistics:

$$\bar{M}_d(q) \equiv \bar{\sigma}_c^2(q) - \bar{\sigma}_a^2$$

$$\bar{M}_r(q) \equiv \frac{\bar{\sigma}_c^2(q)}{\bar{\sigma}_a^2} - 1$$

Although this does not imply an unbiased variance ratio, simulation experiments show that the finite-sample properties of the test statistics are closer to their asymptotic counterparts when this bias adjustment is made. Inference for the overlapping variance differences and ratios may then be performed using Theorem 2.

Theorem 2. Under the null hypothesis H, the asymptotic distributions of the statistics $M_d(q)$, $M_r(q)$, $M_d(q)$, and $M_r(q)$ are given by

$$\sqrt{nq}M_d(q) \sim \sqrt{nq}\bar{M}_d(q) \sim N(0, \frac{2(q-1)(q-1)}{3q}\sigma_0^4)$$

$$\sqrt{nq}M_r(q) \sim \sqrt{nq}\bar{M}_r(q) \sim N(0, \frac{2(q-1)(q-1)}{3q})$$

To develop some intuition for these variance ratios, observe that for an aggregation value q of 2, the M_r,(q) statistic may be reexpressed as

$$M_r(2) = \hat{p}(1) - \frac{1}{4n\hat{\sigma}_a^2}[(X_1 - X_0 - \hat{\mu})^2 + (X_{2n} - X_{2n-1} - \hat{\mu})^2] \cong \hat{p}(1)$$

Hence, for q = 2 the M_r(q) statistic is approximately the first-order autocorrelation coefficient estimator $\hat{p}(1)$ of the differences. More generally, it may be shown that

$$M_r(q) \cong \frac{2(q-1)}{q}\hat{p}(1) + \frac{2(q-2)}{q}\hat{p}(2) + \cdots + \frac{2}{q}\hat{p}(q-1)$$

Where $\hat{p}(k)$ denotes the k_{th} order autocorrelation coefficient estimator of the first differences of X_t. The last equation provides a simple interpretation for the variance ratios computed with an aggregation value q: They are (approximately) linear combinations of the first q – 1 autocorrelation coefficient estimators of the first differences with arithmetically declining weights.

4.5.1. Heteroscedastic increments

Due to raising consensus among financial economists that volatilities actually shows changes over time, a rejection of the random walk hypothesis because of heteroscedasticity would not be a significant achievement. I therefore wish to derive a version of specification test of the random walk model that is robust to changing variances. As long as the increments are uncorrelated, even in the presence of heteroscedasticity the variance ratio should still approach unity as the number of observations increase without bound. It happens because the variance of the sum of uncorrelated increments must still equal the sum of the variances. However, the asymptotic variance of the variance ratios will obviously depend on the type and degree of existing heteroscedasticity. One possible approach is to assume some specific form of heteroscedasticity and then to calculate the asymptotic variance of $\overline{M}_r(q)$ under this null hypothesis. However, to allow for more general forms of heteroscedasticity, I employ an approach developed by White (1980) and by White and Domowitz (1984). This approach also allows to relax the requirement of gaussian increments, an especially important extension in view of stock returns' well-documented empirical departures from normality. Specifically, I consider the null hypothesis H*:

1. For all t, $E(\varepsilon_t) = 0$, and $E(\varepsilon_t \varepsilon_{t-\tau}) = 0$ for any $\tau \neq 0$;
2. ε_t is φ-mixing with coefficients φ(m) of size r/(2r-1) ot is α-mixing with coefficients α(m) of size r/(r-1), where r>1, such that for all t and for any τ ≥0, there exists some δ>0 for which

$$E|\varepsilon_t \varepsilon_{t-\tau}|^{2(r+\delta)} < \Delta < \infty$$

3. $\lim_{nq\to\infty} \frac{1}{nq} \sum_{t=1}^{nq} E(\varepsilon_t^2) = \sigma_0^2 < \infty$
4. For all t, $E(\varepsilon_t \varepsilon_{t-j}, \varepsilon_t \varepsilon_{t-k}) = 0$ for any nonzero j and k where $j \neq k$.

This null hypothesis assumes that X_t, possesses uncorrelated increments but allows for quite general forms of heteroscedasticity, including deterministic changes in the variance (due, for example, to seasonal factors) and Engle (1982) ARCH processes (in which the conditional variance depends on past information).

Since $\overline{M}_r(q)$ still approaches zero under H*, I need only compute its asymptotic variance or $\theta(q)$ to perform the standard inferences. I do this in two steps. First, recall that the following equality obtains asymptotically:

$$\overline{M_r}(q) = \frac{2(q-j)}{q}\hat{p}(j)$$

Second, note that under H* the autocorrelation coefficient estimators $\hat{p}(j)$ are asymptotically uncorrelated. If we can obtain asymptotic variances $\delta(j)$ for each of the $\hat{p}(j)$ under H *, I may readily calculate the asymptotic variance $\theta(q)$ of $\overline{M_r}(q)$ as the weighted sum of the $\delta(j)$, where the weights are simply the weights in previous relation squares In other words in can be expressed as:

Theorem 3. Denote by $\delta(j)$ and $\theta(q)$ the asymptotic variances of $\hat{p}(j)$ and $\overline{M_r}(q)$, respectively. Then under the null hypothesis H*:

1. The statistics $J_d(q)$, $J_r(q)$, $M_d(q)$, $M_r(q)$, $\overline{M_d}(q)$, and $\overline{M_r}(q)$ all converge almost surely to zero for all q as n increases without bound.
2. The following is a heteroscedasticity-consistent estimator of $\delta(j)$:

$$\hat{\delta}(j) = \frac{\sum_{k=j+1}^{nq}(X_k - X_{k-1} - \hat{\mu})^2\left(X_{k-j} - X_{k-j-1} - \hat{\mu}\right)^2}{\left[\sum_{k=1}^{nq}(X_k - X_{k-1} - \hat{\mu})^2\right]^2}$$

3. The following is a heteroscedasticity-consistent estimator o $\theta(q)$f:

$$\hat{\theta}(q) = \sum_{j=1}^{q-1}\left[\frac{2(q-j)}{q}\right]^2\hat{\delta}(j)$$

Despite the presence of general heteroscedasticity, the standardized test statistic

$$z*(q) \equiv \frac{\sqrt{nq}\overline{M}_r(q)}{\sqrt{\hat{\theta}}}$$

is still asymptotically standard normal. In next section I use the $z*(q)$ statistic to test empirically for random walks in weekly stock returns data.

4.5.2. Testing the random walk hypothesis for weekly returns of S&P500

For testing the random walk hypothesis I’ve chosen weekly returns of S&P 500 for the same period which I’d taken in previous test. This is from 01.1994 to 12.2016. The whole population consists of 1199 observations of equal intervals. Weekly sampling is the ideal compromise, providing a large number of observations and minimizing the biases inherent in daily data, that's the reason why I use it.

Table 1 – Market index results for a one-week base observation period

Time period	Number of observations	Number q of base observations aggregated to form variance ratio			
		2	4	8	16
01.1994 - 12.2016	1199	0,91	0,89	0,87	0,95

The variance ratios 1 + $\overline{M_r}(q)$ are reported in the main row. Under the random walk null hypothesis, the value of the variance ratio is 1. Random walk hypothesis cannot be rejected in this case because the M coefficient is below 1 in each q.

5. Conclusion

The question of how generate stock prices are generated has always been important for the researchers. Various hypotheses were created, some of which claimed that prices move randomly. Others have stated that it's possible to identify recurring patterns. This paper was aimed on reviewing different pricing models for stocks, their characteristics and properties.

Different characteristics of time series might divide them into stationary and non-stationary. It is crucial to evaluate them before making hypothesis testing, because the violation of one (or several) conditions of the stationary can cause an error in the calculation of test statistics and consequently, hypothesis testing.

Constant expected return (CER) implies that expected returns are independently and identically distributed, have constant mean and variance. It's a static model, meaning that dependent variable is calculated only by the regressors from the same time period. Markov switching model, in its turn, states that the time series can exist in several states (or regimes). The random behavior of the state variable is controlled by transition matrix.

Many researchers (and investors) hold opposing views on the market stock pricing. Some of them are adherents of the technical analysis approach, other prefer fundamental analysis. The difference between them lies in the source of information on prices. Technical analysis involves the study of past asset prices. Fundamental analyzes the information about the company, news and rumors. Both of them are trying to find some kind of "true value", knowing that may allow to make an extraordinary profit, exceeding the one that will get usual market participants. However, if the random walk of stock prices is a reasonable data generating process, then calculating a "true value" seems pointless. One another possible explainer of changes in securities prices is the social cognitive theory, claiming that people base their actions watching the behavior of others. In the case of the success of others, the individual can upgrade his/her own behavior. This also applies to stock trading process. Therefore, sometimes the market can show seemingly irrational movement, which is impossible to explain by anything except the social cognitive theory approach.

The results obtained in the current study do not allow me to identify unambiguous patterns in the change in share prices. I also cannot unreservedly reject the null hypothesis about random walk in price movement. I've obtained as signs of both the randomness of the volatility of stock prices, and certain repetitive patterns.

I can say with confidence that this topic requires additional deep study and elaboration. Perhaps in the future, new ways of testing will be developed, which will give an accurate answer to the question posed. At the moment it is not possible to say whether changes in stock prices are absolutely random or not. However, as far as I can see, there is no consensus in existing studies on this question either. So, another option is that it is not possible at all to reveal the true nature of stock prices movements.

References

Abarbanell, J. S., & Bushee, B. J. (1997). Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research, 35*(1), 1-24.

Abraham, A., Seyyed, F. J., & Alsakran, S. A. (2002). Testing the random walk behavior and efficiency of the Gulf stock markets. *Financial Review, 37*(3), 469-480.

Araújo, R. d. A., Oliveira, A. L., & Meira, S. (2015). A hybrid model for high-frequency stock market forecasting. *Expert Systems with Applications, 42*(8), 4081-4096.

Asprem, M. (1989). Stock prices, asset portfolios and macroeconomic variables in ten European countries. *Journal of Banking & Finance, 13*(4-5), 589-612.

Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics, 91*(2), 299-323.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 47-78.

Bandura, A. (1997). Self-efficacy: The exercise of control: New York: Freeman.

Barber, B. M., & Odean, T. (1999). The courage of misguided convictions. *Financial analysts journal, 55*(6), 41-55.

Beber, A., & Pagano, M. (2013). Short-selling bans around the world: Evidence from the 2007–09 crisis. *The Journal of Finance, 68*(1), 343-381.

Blume, L., Easley, D., & O'hara, M. (1994). Market statistics and technical analysis: The role of volume. *The Journal of Finance, 49*(1), 153-181.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science, 2*(1), 1-8.

Borges, M. R. (2011). Random walk tests for the Lisbon stock market. *Applied Economics, 43*(5), 631-639.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance, 47*(5), 1731-1764.

Burton, F. E. T., & Shah, S. N. (2017). Efficient Market Hypothesis. *CMT Level I 2017: An Introduction to Technical Analysis*.

Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance, 43*(3), 661-676.

Chancharat, S., & Valadkhani, A. (2007). Testing for the random walk hypothesis and structural breaks in International stock prices.

Charles, A., & Darné, O. (2009). The random walk hypothesis for Chinese stock markets: Evidence from variance ratio tests. *Economic Systems, 33*(2), 117-126.

Chaudhuri, K., & Wu, Y. (2003). Random walk versus breaking trend in stock prices: Evidence from emerging markets. *Journal of Banking & Finance, 27*(4), 575-592.

DiFonzo, N., & Bordia, P. (2002). Rumors and stable-cause attribution in prediction and behavior. *Organizational Behavior and Human Decision Processes, 88*(2), 785-800.

Dougherty, C. (2007). *Introduction to econometrics*: Oxford University Press, USA.

Dsouza, J. J., & Mallikarjunappa, T. (2015). Does the Indian Stock Market Exhibit Random Walk? *Paradigm, 19*(1), 1-20.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Fama, E. F. (1995). Random walks in stock market prices. *Financial analysts journal, 51*(1), 75-80.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial Economics, 22*(1), 3-25.

Godfrey, M. D., Granger, C. W., & Morgenstern, O. (1964). The random-walk hypothesis of stock market behavior. *Kyklos, 17*(1), 1-30.

Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications, 38*(8), 10389-10397.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, 1251-1271.

Hirshleifer, D., Subrahmanyam, A., & Titman, S. (2006). Feedback and the success of irrational investors. *Journal of financial Economics, 81*(2), 311-338.

Hui, E. C., & Yam, S.-C. P. (2014). Can we beat the “buy-and-hold” strategy? Analysis on European and American securitized real estate indices. *International Journal of Strategic Property Management, 18*(1), 28-37.

Jarrett, J. E. (2008). Random walk, capital market efficiency and predicting stock returns for Hong Kong Exchanges and Clearing Limited. *Management Research News, 31*(2), 142-148.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of financial Economics, 17*(2), 357-390.

Kilian, L., & Taylor, M. P. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics, 60*(1), 85-107.

Kuan, C.-M. (2002). Lecture on the Markov switching model. *Institute of Economics Academia Sinica*, 1-30.

Lean, H. H., Mishra, V., & Smyth, R. (2015). The relevance of heteroskedas-ticity and structural breaks when testing for a random walk with high-frequency financial data: Evidence from asean stock markets. *The Handbook of High Frequency Trading*, 59-73.

Lee, C.-C., Lee, J.-D., & Lee, C.-C. (2010). Stock prices and the efficient market hypothesis: Evidence from a panel stationary test with structural breaks. *Japan and the world economy, 22*(1), 49-58.

Lesmond, D. A., Schill, M. J., & Zhou, C. (2004). The illusory nature of momentum profits. *Journal of financial Economics, 71*(2), 349-380.

Lim, K.-P., & Brooks, R. D. (2010). Why do emerging stock markets experience more persistent price deviations from a random walk over time? A country-level analysis. *Macroeconomic Dynamics, 14*(S1), 3-41.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies, 1*(1), 41-66.

Lock, D. B. (2007). The Taiwan stock market does follow a random walk. *Economics Bulletin, 7*(3), 1-8.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspectives, 17*(1), 59-82.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417.

Mishra, A., Mishra, V., & Smyth, R. (2015). The random-walk hypothesis on the Indian stock market. *Emerging Markets Finance and Trade, 51*(5), 879-892.

Moosa, I., & Vaz, J. (2015). Why is it so difficult to outperform the random walk? An application of the Meese–Rogoff puzzle to stock prices. *Applied Economics, 47*(4), 398-407.

Nalewaik, J. (2011). Forecasting recessions using stall speeds.

Neftci, S. N. (1991). Naive trading rules in financial markets and wiener-kolmogorov prediction theory: a study of" technical analysis". *Journal of Business*, 549-571.

Okpara, G. C. (2010). Stock market prices and the random walk hypothesis: Further evidence from Nigeria. *Journal of Economics and International Finance, 2*(4), 49.

Oskooe, S. A. P. (2011). *The Random Walk Hypothesis in Emerging Stock Market-Evidence from Nonlinear Fourier Unit Root Test.* Paper presented at the Proceedings of the World Congress on Engineering.

Poshakwale, S. (1996). Evidence on weak form efficiency and day of the week effect in the Indian stock market. *Finance India, 10*(3), 605-616.

Rahman, A., & Saadi, S. (2008). Random walk and breaking trend in financial series: An econometric critique of unit root tests. *Review of Financial Economics, 17*(3), 204-212.

Shiller, R. J., & Perron, P. (1985). Testing the random walk hypothesis: Power versus frequency of observation. *Economics Letters, 18*(4), 381-386.

Solnik, B. H. (1973). Note on the validity of the random walk for European stock prices. *The Journal of Finance, 28*(5), 1151-1159.

Spinu, F. (2015). Buy-and-hold versus constantly rebalanced portfolios: A theoretical comparison. *Journal of Asset Management, 16*(2), 79-84.

Suleman, M. T., Hamid, K., Ali Shah, S. Z., Akkash, I., & Shahid, R. (2010). Testing the weak form of efficient market hypothesis: Empirical evidence from Asia-Pacific markets.

Taylor, M. P., & Allen, H. (1992). The use of technical analysis in the foreign exchange market. *Journal of international Money and Finance, 11*(3), 304-314.

Urrutia, J. L. (1995). Tests of random walk and market efficiency for Latin American emerging equity markets. *Journal of financial research, 18*(3), 299-309.

Wang, F., & Landau, D. (2001). Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram. *Physical Review E, 64*(5), 056101.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838.

White, H., & Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica: Journal of the Econometric Society*, 143-161.